

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Ingegneria Biomedica, Elettrica e dei Sistemi

Ciclo 33°

**Settore Concorsuale:** 09/G2 - Bioingegneria

**Settore Scientifico Disciplinare:** ING-INF/06 – Bioingegneria Elettronica e Informatica

METHODS FOR ACQUISITION AND INTEGRATION OF PERSONAL  
WELLNESS PARAMETERS

**Presentata da:** Veronica Chiara Zuccalà

**Coordinatore Dottorato**

Michele Monaci

**Supervisore**

Giuseppe Coppini

**Co-supervisori**

Stefano Diciotti

Mauro Ursino

**Esame finale anno 2021**

# Abstract

Wellness indicates the state or condition of being in good physical and mental health. Stress is a common state of emotional strain that plays a crucial role in the everyday quality of life, and significantly affects the wellness state of a person, being a ubiquitous risk factor for virtually all non-communicable diseases. Nowadays, there is a growing individual awareness of the importance of a proper lifestyle and a generalized trend to become an active part in monitoring, preserving, and improving personal wellness for both physical and emotional aspects. In this respect, it must be pointed out that the majority studies in this field relies on the evaluation of the changes of sensed parameters passing from rest to “maximal” (intense) stress. However, the vast majority of people usually experiences stressing circumstances in everyday life (e.g. at home, at work) which are prompted by a wide spectrum of stimuli having varying intensity and consciousness. This led us to investigate the impact of mild cognitive activation which can be somehow comparable to usual situations that everyone can face in daily life.

As documented in scientific literature, several signals and data can be useful to characterize the state of a person, but not all of them are equally important and/or reliable. In order to reduce the complexity of data acquisition procedures and to simplify the modelling of individual wellness, it is crucial to analyse the mutual relevance of the different pieces of information. In this work we focus on a subset of well-established psychophysical descriptors, including heart rate and heart rate variability, respiratory rate, electrodermal activity, and electrical brain activity.

A deepened investigation was carried out to identify a set of devices enabling the measurement of the psycho-physical parameters mentioned above. Beside technical and methodological constraints on quality of collected data, the design of the experimental setup and the selection of sensing devices was also driven by qualitative criteria such as intrusiveness, reliability, and ease of use. These are deemed crucial for implementing effective (self-)monitoring strategies.

During the PhD work, a reference dataset, named “Mild Cognitive Activation” (MCA), was collected. It includes signals and data from a group of volunteers according to a protocol approved by the CNR Committee for Research Ethics and Bioethics. Data collection was focused on the impact of a mild cognitive activation induced by a simple test.

The last aim of the project was the definition of a quantitative model for data integration providing a concise description of the wellness status of a person. This process was based on unsupervised learning paradigms. In this phase of the work, data from MCA were integrated with data from the “Stress Recognition in Automobile Drivers” (SRAD) dataset (MIT Media Lab). This allowed a cross validation of the integration methodology.

# Contents

Abstract.....	2
List of figures.....	3
List of tables .....	5
Acronyms .....	6
Chapter 1. Personal wellness.....	8
1.1. Objective vs subjective wellness scales .....	8
1.2. Wellness and Self-Monitoring.....	10
1.3. Wellness sensing.....	13
1.4. A ubiquitous wellness factor: stress level.....	14
1.5. Targets of the project.....	15
Chapter 2. Signals and sensors .....	17
2.1. Psycho-physical parameters .....	17
2.2. Experimental setup .....	20
2.3. Testing protocol .....	23
Chapter 3. Heart rate and heart rate variability from single-channel video .....	25
3.1. Introduction .....	25
3.2. Blood volume pulse from video .....	27
3.3. Materials and methods .....	28
3.3.1. Study's participants.....	28
3.3.2. Experimental setup and video acquisition.....	28
3.3.3. ECG signal analysis .....	29
3.3.4. ROI's selection.....	30
3.3.5. Video signal analysis .....	30
3.3.6. Recording length .....	31
3.3.7. Video frame rate .....	32
3.4. Results.....	32
3.4.1. Analysing video at the native frame rate.....	32
3.4.2. Varying the frame rates .....	36
3.4.3. Varying the acquisition time .....	38

3.5.	Discussion.....	40
Chapter 4.	Unobtrusive monitoring of stress indicators at rest .....	42
4.1.	Materials and methods .....	42
4.1.1.	Study's participants and experimental setup .....	42
4.1.2.	Data analysis .....	42
4.2.	Results.....	43
4.3.	Discussion.....	44
Chapter 5.	Unobtrusive monitoring of EEG signals during mild cognitive activation.....	45
5.1.	Materials and methods .....	45
5.1.1.	Study's participants and experimental setup .....	45
5.1.2.	Data analysis .....	45
5.2.	Results.....	47
5.3.	Discussion.....	47
Chapter 6.	Unobtrusive monitoring of stress indicators during mild cognitive activation .....	49
6.1.	Materials and methods .....	49
6.1.1.	Study's participants and experimental setup .....	49
6.1.2.	Data analysis .....	49
6.2.	Results.....	50
6.3.	Discussion.....	52
Chapter 7.	Data integration .....	53
7.1.	Datasets .....	53
7.1.1.	The MCA dataset description and its analysis .....	53
7.1.2.	The SRAD dataset .....	53
7.1.3.	The SRAD dataset analysis .....	54
7.2.	Features selection .....	56
7.2.1.	Sequential forward feature selection .....	56
7.2.2.	Auto-Encoder neural network .....	57
7.3.	Detection of activation status.....	58
7.3.1.	Activation representation by k-means clustering.....	58
7.3.2.	Activation representation by SOM maps.....	58
7.4.	Results.....	59
7.4.1.	Sequential forward feature selection .....	59
7.4.2.	Auto-Encoder neural network .....	62
7.4.3.	K-means clustering.....	63

7.4.4. Self-Organizing Maps .....	64
7.5. Discussion.....	67
Conclusions .....	69
Appendix A .....	71
Appendix B .....	72
Bibliography .....	79

# List of figures

Figure 1 Mobile phone as mobile health platform. ....	10
Figure 2 An activity-based self-wellness monitoring application. ....	11
Figure 3 A four-layer conceptual model for wellness state monitoring. ....	12
Figure 4 Experimental setup. ....	22
Figure 5 Test A on the left and Test B on the right. ....	24
Figure 6 Data acquisition setup. ....	29
Figure 7 Dimensions and positions of the ROIs (ROI1 on the forehead area, ROI2 on the right cheek, and ROI3 on the left cheek). $d$ is the distance between the centres of the rectangles around the eyes and $l$ is the side of the square that is set to 20 pixel. ....	30
Figure 8 Procedure based on ICA. ....	31
Figure 9 Scatter plots comparing NN and SDNN between the video signal and ECG signal. The video signal data are obtained from $x_1(t)$ applying M1, on the left side, and from $y_s(t)$ applying M2, on the right side. ....	35
Figure 10 Scatter plots comparing LF, HF, and LF/HF values between the video signal and ECG signal. The video signal data are obtained from $x_1(t)$ . ....	36
Figure 11 A typical GSR record. The raw and the filtered signals of the skin conductivity of a subject during the resting state are shown. The regression line (obtained by the linear regression) is the average slope of the signal. ....	43
Figure 12 The EEG signals after ICA processing. ....	46
Figure 13 An example of GSR signal. Upward the raw and the filtered signal, below the phasic signal and the peaks detection. ....	50
Figure 14 Box plot of the NN value of the ECG signal (left) and the video signal (right) in the three experimental conditions. ....	50
Figure 15 The box plot of the RR value in the three experimental conditions. ....	51
Figure 16 Placement of recording instruments as deployed while collecting the “Stress Recognition in Automobile Drivers” dataset. Picture reproduced from [120]. ....	54
Figure 17 Accuracy of SFFS process. On the left: accuracy of LDA classifier evaluated on the SRAD dataset varying the number of feature according to SFFS ranking. On the right: accuracy evaluated on MCA. ....	62
Figure 18 Accuracy of AE process. On the left: accuracy of LDA classifier evaluated on the SRAD dataset with different auto-encoding dimensionality. On the right: accuracy evaluated on MCA. ....	62
Figure 19 Classification performance for the SRAD dataset obtained using the first two features selected by the SFFS method. On the left a scatter plot of the dataset in the feature space, indicating the two clusters and related centroids. On the right the validation of the clusterization process. ....	63
Figure 20 Classification performance for the MCA dataset obtained using the first two features selected by the SFFS method. On the left a scatter plot of the dataset in the feature space, indicating the two clusters and related centroids. On the right the validation of the clusterization process. ....	64
Figure 21 Maps of distances between the weight of neighboring units for 4 x 4 and 5 x 5 maps. Darkest. ....	65
Figure 22 Maps of weight components. Darkest colors indicate smallest values while light colors denotes largest ones. ....	66

Figure 23 Map of winning units for SRAD data according to rest and activation labels. ....	67
Figure 24 Map of winning units for MCA data according to rest and activation labels. ....	67

# List of tables

Table 1 Main psycho-physical parameters investigated in this work. Their major scope, standard (or usual) source, as well data accessibility and reliability especially for unobtrusive (self-)monitoring are outlined.....	17
Table 2 Mean, standard deviation, minimum, and maximum values of the data obtained by ECG and video signal analysis at the native frame rate. ....	33
Table 3 Absolute errors (mean, standard deviation, minimum, and maximum), median of signer errors, Pearson coefficient, and spearman's obtained using M1 and M2, respectively. ....	34
Table 4 Mean, standard deviation, minimum, and maximum values of the data obtained by video signal analysis varying the frame rates. ....	37
Table 5 Absolute errors (mean, standard deviation, minimum, and maximum, median of signed errors, Pearson coefficient, and spearman's obtained at three different video rates for $x_1(t)$ and $y_5(t)$ using M1 and M2, respectively. ....	38
Table 6 Mean, standard deviation, minimum, and maximum values of the data obtained by ECG and video signal analysis varying the acquisition time from M1.....	39
Table 7 Absolute errors (mean, standard deviation, minimum, and maximum), median of signed errors, Pearson coefficients, and spearman's for five different durations of videos for $x_1(t)$ and $y_5(t)$ using M1 and M2. ....	39
Table 8 Results summary. For each group, the median and the interquartile range, in brackets, of HR and HRV parameters (NN, SDNN, LF, HF, LF/HF), respiratory rate, and the slope of the skin conductivity trend ( $GSR_{SLOPE}$ ).....	44
Table 9 For each condition, the median and the interquartile range, in brackets, of the all parameters extracted. ....	47
Table 10 Median peak value ( $\mu S$ ) and number of peaks of the GSR for the three experimental conditions.....	51
Table 11 Extracted features and related source signals. ....	55
Table 12 The score and the selection frequency (SF) obtained by each features. ....	60
Table 13 SRAD database: median RR and number of GSR peaks both at rest and during driving. Interquartile range is reported in brackets.....	60
Table 14 MCA database: median RR with interquartile range (in brackets) together with the number of GSR peaks both at rest and during SCWT. No interquartile range is available for number of GSR peaks as we only have a single block. ....	61



# Acronyms

AE auto-encoder

AE<sub>min</sub> absolute minimum error

AE<sub>max</sub> absolute maximum error

AF antero – frontal

ANS autonomic nervous systems

APA alpha power asymmetry

BPA beta power asymmetry

BSS blind source separation

BVT blood volume pulse

DPA delta power asymmetry

ECG electrocardiography

EDA electrodermal activity

EEG electroencephalography

EMG electromyography

GSR galvanic skin response

HF high frequency

HR heart rate

HRV heart rate variability

IBI inter beat interval

ICA independent component analysis

iPPG imaging photoplethysmography

LDA linear discriminant analysis

LF low frequency

MAE mean absolute error

$M_e$  median of signed errors

MCA mild cognitive activation

NCDs non-communicable diseases

NN normal to normal interval

PCA principal components analysis

PPG photoplethysmography

PSD power spectrum density

RGB read green blue

ROI region of interest

RR respiratory rate

SCWT Stroop colour word test

$SD_{AE}$  absolute errors standard deviation

SDNN standard deviation of the normal to normal interval

SFFS sequential forward feature selection

SOM self-organized map

SRAD stress recognition in automobile drivers

SVM support vector machine

TP temporo - parietal

TPA theta power asymmetry

WM Wize Mirror

# Chapter 1. Personal wellness

Wellness indicates the state or condition of being in good physical and mental health. According to the World Health Organization, health should be defined as a state of complete physical, mental, and social wellness, and not merely as the absence of disease and infirmity [1]. In this perspective, any effort to evaluate personal wellness needs to be framed in a global view encompassing virtually all aspects of individual life. The authors of [2] propose a definition of wellness as a multidimensional state of being possessed by the individual, which is part of health and can be described as positive health; quality of life and a sense of well-being are the descriptors of wellness. In [2] it is pointed out that wellness is not the same as physical fitness, is not a form of alternative medicine, and is what one is rather than what one does. In [3] wellness is defined as a state of well-coordinated, goal-oriented functioning to maximize personal potential and to enhance the quality-of-life.

Critical issues related to wellness include the development of valid and reliable methods for assessing wellness and the study of the factors that impact on wellness once good measures are established. The first attempts to construct numerical indices for wellness evaluation date back to the 1940s. Since the 1960s, more general tools for quantifying individual health and wellness have been proposed, which pay attention to individuals in their many facets, including physiological measurements, psychological traits, and social parameters.

Over the last years, the interest in the role of individual wellness in disease prevention has grown rapidly. The scientific and medical community has started focusing on the need to provide people with knowledge, services, and tools to support a healthy lifestyle. Indeed, our everyday choices, along with our social and economic environment, influence the way our genes express themselves in our dynamic phenotype and contribute to health or disease[4][5]. This is especially relevant for non-communicable diseases (NCDs), such as heart disease and stroke, cancer, diabetes, and chronic lung disease. Worldwide, NCDs are responsible for about two-thirds of deaths. Moreover, the prevalence of NCDs is expected to increase dramatically.

## 1.1. Objective vs subjective wellness scales

The health of an individual can be analysed from different viewpoints and with different objectives. A consolidated literature exists on the measurement of the wellness of a population [6], motivated by economic or epidemiological reasons. The indices are computed via statistics on large population samples, and are usually based on the frequency

and duration of single diseases and related mortality [7]. While the indicators are useful to characterize the population and drive policy actions by health services, they are not suited to model the health status of individuals, drive behavioural changes, and prevent diseases. Hence, we will limit our analysis to the assessment of individual wellness.

Historically, the early attempts to quantify the status of a person were focused on the assessment of the functional abilities of a patient in specific pathophysiological conditions. Noteworthy examples are the American Rheumatism Association Function Scale [8], the Karnofsky Score [9] created for cancer patients, and the New York Heart Association Functional Classes [10]. These simple instruments were conceived to combine several dimensions into a single scale. They also served to define a common and consistent terminology in the medical community. These tools are based on the judgment of an external observer, typically a physician, and are therefore referred to as objective.

By contrast, subjective scales are based on subjective perceived wellness. A relatively simple approach to gather information about personal health information is the use of questionnaires or remote consultation as forms participative inquiry [11]. The International Wellbeing Group [12] proposes a questionnaire-based method to compute a personal Wellbeing Index, based on a domain-level representation of global life satisfaction: individual items refer to specific life domains, and the scores are averaged to produce a measure of Subjective Wellbeing. A well-known tool for assessing the perceived health status is the 36-item Short Form questionnaire that measures eight constructs of functioning: physical function, mental health, role emotional, role physical, social function, bodily pain, vitality, and general health [13]. The study [14] tries to combine the relative contributions of both physical and psychological health in a measure of global well-being. The study is carried out in a sample of healthy young adults, early in their college careers, with the purpose of exploring individual differences that contribute to the sense of well-being, and identifying characteristics of flourishing versus non-flourishing individuals. Along with age and sex, the study uses indices derived from the following questionnaires: Arizona Integrative Outcomes Scale, global physical health, positive and negative mood (Positive and Negative Affect Schedule or PANAS), resilience (Connor-Davidson Resilience Scale), and repressive defensiveness (Marlowe-Crowne Social Desirability Scale or MCSD-SF). The results are based on the ratio of positive/negative scores established by using complex systems methods. According to the study results, positive-to-negative affect (P/N) ratio explains a substantial portion of the variance in the well-being of healthy young adults and represents a well-being measure easy to compute, sensitive, and useful as a clinical measure to track change with treatment over time.

Despite the number of subjective measures reported in the literature, there is substantial evidence about the discomfort that traditional information gathering methods produce [15]: people are asked to fill lengthy questionnaires, answer a huge set of questions, and provide

a large amount of personal information about their health status. On the other hand, individual perception of personal wellness is a relevant piece of information about health status and should be taken into account when defining/measuring the status of an individual. It stands to reason that integrating objective measurements with self-reported data should provide the basis for comprehensive and practicable wellness indices.

## 1.2. Wellness and Self-Monitoring

There is a growing individual awareness of the importance of a proper lifestyle and a generalized trend to become an active part in monitoring, preserving, and improving personal wellness. The self-monitoring and the self-learning systems are important tools to provide the knowledge, the customer care, and the devices for the achievement of the personal wellness. Recently, the use of mobile devices has opened a new perspective on the assessment of the wellness and stimulated the development of new approaches based on individual self-monitoring. In recent years, a growing number of studies addressed the development of systems for quantifying the individual health and wellness status [16].

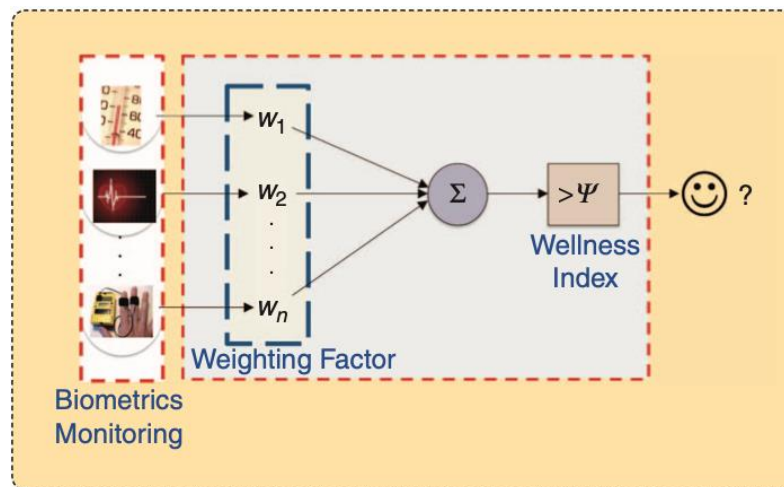


Figure 1 Mobile phone as mobile health platform.

The 2010 position paper by Kailas [17] provides a synthetic view on the paradigm of wellness mobiles, wherein mobile phones are equipped with biometric sensors for real-time self-wellness monitoring. A possible platform could follow the scheme in Figure 1. Inexpensive sensors measure and record biometrics of the user, such as temperature, heart rate (HR), and galvanic skin response (GSR). The biometrics are then weighted (weighting coefficients denoted as  $w_1, \dots, w_n$ , in Figure 1 and combined with other information, possibly also obtained using the mobile phone. The result is a wellness index that can be customized to the users' daily lifestyle. In this way, the concept of wellness mobile can improve the quality of life of the phone user by facilitating timely and better-quality measurements,

supporting instantaneous feedback, improving the quality of medical information, and enhancing patient compliance.

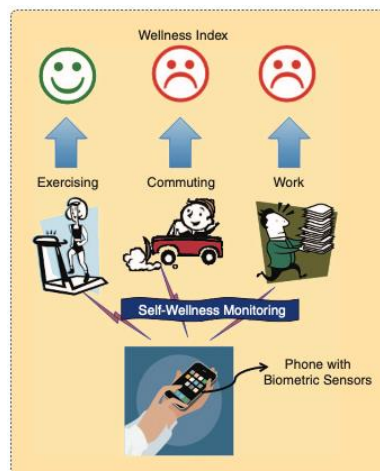


Figure 2 An activity-based self-wellness monitoring application.

Further, the mobile application can be used during various activities (Figure 2), which entail different demands from the user to track the wellness states transitions during different tasks.

The statements in the position paper [17] are reinforced in [18], which sketches the guidelines for defining a wellness inference algorithm based on wellness recognition and tracking, followed by user assistance. As shown in Figure 2, a conceptual model for affective state monitoring on a mobile wellness platform should consist of four layers (Figure 3):

1. Predictive Layer (Activities). The layer includes the factors that affect the wellness state of a user, under the hypothesis that the state of wellness undergoes variations during the day according to different activities carried out: user profile, context, time of the activity, and prior wellness state.
2. Hidden Layer (Wellness State). Well-known wellness inhibitors are stress, fatigue, bad health, and confusion, which can cause various diseases or detectable conditions such as large deviations from normal processes and biometrics (e.g., body temperature, heart rate, etc.). The model recognizes and tracks the anti-wellness (e.g., stress) state transitions using sensory measurements. Each anti-wellness state can have a set of possible values.
3. Observable Layer (Evidence). An evidence is an observable feature that is capable of providing clues about the users' internal wellness state that is hidden. There are four classes of measurable evidence:

- a. the user's physical appearance features: visual features that characterize eyelid and pupil movement (gaze tracking, pupillary response, etc.), gestures, change in pressure at finger tips, and head movement;
  - b. physiological measures: electrocardiography (ECG) measures, electromyography (EMG) measures, galvanic skin response measures that assess the electrical properties of the skin in response to different kinds of stimuli, etc.;
  - c. behavioural traits, such as key strokes using the keyboard;
  - d. performance measures, as user response time and accuracy rate in a task, which can be influenced by wellness states.
4. Action Layer (Timely User Assistance). The design of an appropriate assistance depends on a number of factors such as the activity, and context, and should be personalized to the user. The range of applications and assistive services offered to the user could range from advice/tips from medical on-line professionals to entertainment (e.g., on-demand video, music, etc.) to online games.

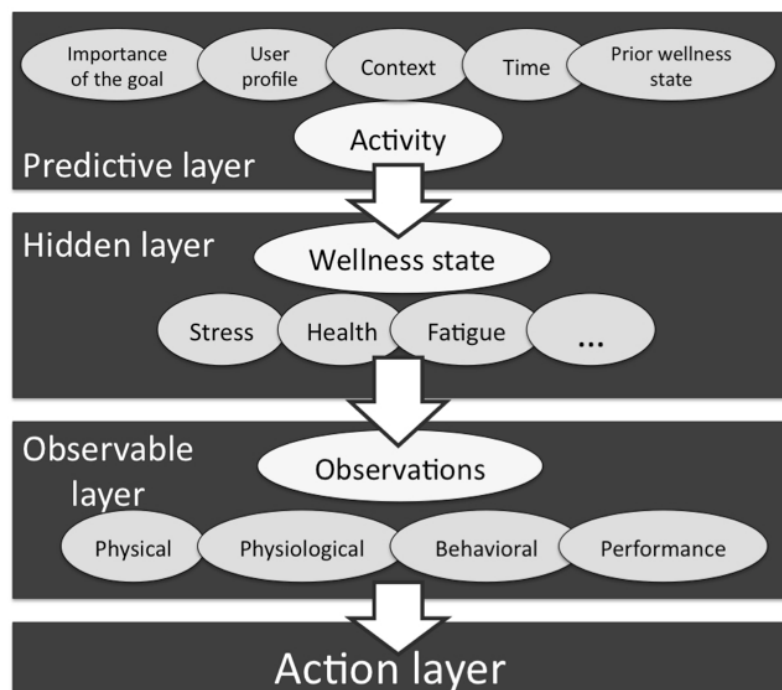


Figure 3 A four-layer conceptual model for wellness state monitoring.

### 1.3. Wellness sensing

The interest in measuring wellness status and related aspects is documented by a huge scientific and technical literature on the topic, exponentially grown in last decade. A detailed survey of published works goes beyond the aims of this thesis. Here we limit to overview the main contributions pertaining the content of next chapters. It should be pointed out that the field is highly interdisciplinary, involving the work of clinical researchers, engineers, physicists with strong contribution from ICT academic community. Addressed topics include investigation and development of:

- a. physical sensors and devices to improve acquisition of physiological signals;
- b. methods and tools for data processing and analysis;
- c. models for feature selection and integration.

A major issue underlying the whole research field pertains the operation in daily-life setting, possibly in naturalistic or quasi-naturalistic conditions, minimizing sensing obtrusiveness. This makes data acquisition extremely prone to different sources of artefacts such user motion, environmental interferences, giving rise to an additional de facto research line aiming to reduce artefact occurrence and their impact on data quality.

In this view, there are many papers dealing with use of smart phone to track minor vital signs and physical activity. Huang et al. [19], used the camera of a smartphone to measure the heart rate and its variability. These parameters could be measured through t-shirts provided by textile sensors in order to monitor the state of elderly people [20]. The sensorized t-shirts were integrated with Wi-Fi, RFID, and GSP/GSM technologies to obtain information about the subject position also and to transmit all the information captured to specialists. In this way, if abnormal parameters have been measured it is possible to intervene quickly. Some studies describe the integration of the psycho-physical parameters to provide information about the individual state such as the fatigue [21][22][23][24]. Al Libawy et al. [21] developed a fitness watch that collect data such as heart rate, temperature, skin conductivity, and provide information on the fatigue of six subjects. In this study, two different classification methods are compared: the Support vector machine (SVM) and the feed forward neural networks. Tayibnapis et al. [23], through imaging techniques analysed the gaze direction, the eye blinking, the frequency of yawns, the heart rate, the heart rate variability (HRV), and the respiratory rate (RR). All these pieces of information were used to estimate the level of fatigue in a driver. The level was estimated through an SVM system.



The studies previously cited, used different psycho-physical parameters to obtain information on the wellness state of a person. Indeed, the psycho-physical parameters usually involved in the description of the state of a person are numerous such as heart rate, heart rate variability, respiratory rate, blood pressure, body temperature, skin conductivity, and electroencephalography signal, but not all of these are equally relevant and reliable. In order to reduce the complexity of the acquisitions and to simplify modelling tasks, it is important to investigate the relevance of the various parameters [25]. Furthermore, it is important to define the relationship between the parameters and the individual condition. For example, the heart rate increases due to both exercise and a fright.

Nowadays, research is moving towards development of integrated multi-sensing systems able to trace and integrate psycho-physical parameters [26][27][28][29][30]. A significant example is provided by the Wize Mirror (WM), a multisensory platform implemented as a smart mirror [30]. The Wize Mirror detects and monitors over time semiotic face signs related to cardio-metabolic risk, and encourages users to reduce their risk by improving their lifestyle. The WM has a user-friendly interface and different devices that measured psycho-physical parameters of the user in front of the mirror. The WM collects multispectral images, video, 3D scans, and information on the molecules in the breath samples (through an external device called Wize Sniffer that contains gas sensors). All the parameters extracted are integrated in order to define a virtual model of the person and a wellness index is calculated. The wellness index provides to the user an auto assessment of the his/her wellness state over time.

To conclude this section, we would like to observe that, despite the increasing interest on the matter, no general (self-)monitoring framework has been established for assessment of the an individual's wellness status by mobile devices. Similarly, no general indicator of the individual health status has gained universal acceptance up to now. In this respect, many authors agree that AI-based methods able to capture complex knowledge representations from data, could provide a valid tool for modelling the individual's state of wellness [31][32].

#### **1.4. A ubiquitous wellness factor: stress level**

Stress is a common state of emotional strain that plays a crucial role in everyday quality of life, so as to the wellness state of a person. This state consists of several complementary and interacting components (i.e. cognitive, affective, and psycho-physiological). Stress is usually a state of tension that is created when a person responds to demands and pressures arising from external sources (e.g. work, family, and social environment) as well as those internally generated (e.g. self-imposed demands and obligations, self-criticism) [33]. Furthermore, chronic stress carries a wide range of health-related diseases, including cardiovascular diseases, cerebrovascular diseases, diabetes, and immune deficiencies [34].

Stress is known to induce abnormal responses in the autonomic nervous system (ANS), which consists of the sympathetic nervous system and the parasympathetic nervous system under antagonistic control [35][36]. These two systems are related to stress and relaxation reactions, respectively, so that stress activates the sympathetic nervous system and suppresses the parasympathetic nervous system [37][38].

Due to the adverse effects of stress in our daily life, stress management has been receiving an increasing attention in health-care and wellness research [39]. As a matter of fact, stress is recognized as a major risk factor for most NCDs and its evaluation is crucial for defining individual wellness.

The majority of studies in this research field relies on the evaluation of the changes of sensed parameters passing from rest to “maximal” (intense) stress. This provides important cues about the individual capability to react to severe stress. On the other hand, the vast majority of people usually experiences stressing stimuli in everyday life (both family and at-work) and a wide spectrum of stressing circumstances, both positive and negative, can be listed.

This led us to investigate the impact of mild stimulations which can be somehow comparable to a normal condition that everyone can deal with in a daily life. This approach also eases monitoring individual status in routine setting (e.g. at work) making it possible to design minimally obtrusive monitoring/testing procedures.

## **1.5. Targets of the project**

In this work we aimed at investigating the main aspects of individual psycho-physical wellness based on sensing self-monitoring. The scope is mainly the every-day life setting with emphasis on stressing conditions.

The project has multiple aims, including:

1. investigation about which are the most relevant psycho-physical parameters involved in the description of the wellness state of a person (Chapter 2);
2. definition of a set of unobtrusive devices for measuring such parameters (Chapter 2, Chapter 3);
3. definition of the experimental setup (Chapter 2, Chapter 3);
4. collection of data from a group of volunteers during a rest state and a mild cognitive activation (Chapter 4, Chapter 5, Chapter 6);

5. definition of a data integration model that could be used to automatically monitor psycho-physical parameters (Chapter 7).

## Chapter 2. Signals and sensors

As already pointed out in Chapter 1, the signals and data useful to characterize the state of a person are numerous, but not all of them are equally important and/or reliable. In order to reduce the complexity of acquisition procedures and to simplify modelling tasks, it is crucial to analyse the mutual relevance of the different pieces of information [25]. In this work we focus on a subset of well-established psychophysical descriptors including: heart rate and its variability, respiratory rate, electrodermal activity, and electrical brain activity. These are listed in Table 1, where we summarize their main properties in view of their usage for wellness monitoring in daily life settings.

Table 1 Main psycho-physical parameters investigated in this work. Their major scope, standard (or usual) source, as well data accessibility and reliability especially for unobtrusive (self-)monitoring are outlined.

Parameter	Relevant to	Data Source	Accessibility	Reliability
Heart Rate	Homeostasis, Emotional status	PPG, ECG	High	High
Heart Rate Variability	Homeostasis, ANS activity	ECG	High/Medium	High/Medium
Respiratory Rate	Homeostasis, Emotional status, General health status	Chest Cage motion	High	High
Electrodermal activity	ANS activity, Emotional status	GSR signal	High	Medium
Electrical Brain activity	In this work we focus on Emotional status, cognitive activity	EEG	Medium/Low	Medium

### 2.1. Psycho-physical parameters

As described in [39], the autonomic nervous system controls the organs of our body such as the heart, the stomach, and the gut. ANS includes sympathetic and parasympathetic nervous systems. The parasympathetic nervous system is responsible for nourishing, calming the nerves to return to the regular function, healing, and regeneration. On the

contrary, the sympathetic system is accountable for activating the glands and the organs for defending the body from the threat. The activation of the sympathetic nervous system might be accompanied by many physical reactions, such as an increase in the heart rate, rapid blood flow to the muscle, activation of sweat glands (with related changes of electrodermal activity - EDA), and increase of respiratory rate.

Heart rate and heart rate variability are significant indicators of the psychophysical status of an individual and are useful clues for detecting risky conditions. HR is the mean number of the heart beats per minute. It is a basic indicator of the cardiovascular homeostasis. The HR varies according to the body's physical needs, changes being observed in a variety of conditions including physical exercise, sleep, anxiety, stress, illness, and assumption of drugs. Monitoring the heart rate is therefore important in both normal and disease conditions. In illness an association exists between HR and outcome in heart failure and high baseline heart rate is considered a cardiovascular risk factor [40]. The HRV is the fluctuation in the time intervals between adjacent heartbeats. It is an index of the adaptation of the heart to circumstances by detecting and readily responding to unpredictable stimuli. The HRV is mainly modulated by the sympathetic and parasympathetic components of the autonomic nervous system [41]. In particular, the sympathetic stimulation is activated in response to stress, exercise, and heart disease and this causes an HR increase [41]. Parasympathetic activity is the result of the function of internal organs, reaction to trauma, allergic reactions, and the inhalation of irritants. This activity determines a decrease of HR [41]. The HRV is altered in several cardiac diseases [42]. In addition, studies have also shown that smoking reduces the HRV due to the increase of the sympathetic activity and reduction of the vagal activity [43]. As a matter of facts, HRV is an indicator of health status in the general population [44], of adaptation to stress in athletes [45], and of fatigue in drivers [46]. Furthermore, the HRV is important to measure mental stress and, coupled with the HR, can be used to monitor individual wellness in behavioural research [47].

Respiratory rate provides important information on a person's health condition and physiological stability, an abnormal respiratory rate being a strong indicator of illness [48]. In fact, a sudden change in respiratory rate is one of the strongest predictors of mortality [49][50]. Respiratory rate is correlated with emotional status and can be used for stress and anxiety detection [51]. Studies reports that respiration rate increases significantly under stressful situations [52].

Galvanic skin response is used in relation to mental state, such as stress, drowsiness and engagement [53]. GSR or EDA is the measure of the continuous changes in the electrical characteristics (conductance) of the skin caused by the variation of the sweating of the human body. This concept is based on the assumption that skin resistance varies with the state of the sweat glands in the skin (the resistance was just reciprocated to determine the conductance). This concept is based on the activity of the ANS as a (strong) stimulation of its

sympathetic branch activates the sweat glands. This tends to decrease the skin resistance, increasing skin conductance [53]. Skin conductivity is sensitive to many different stimuli (strong emotion, a startling event, pain, exercise, deep breathing, a demanding task, cognitive workload, stress etc.); thus, it is often hard to determine what caused a particular skin conductivity response. Different studies noted that the electrodermal response represents an adequate measure for sympathetic activation that is related to stress [54].

The Electroencephalogram (EEG), being a major manifestation of brain activity, is a rich source of information important for detecting and assessing mental stress [55][56][57]. Neurophysiological studies [58] have reported the relationship between human emotion and hemispheric specialization, where the left hemisphere is more involved in processing positive emotions and approach-related behaviours, and the right hemisphere is more involved in processing negative emotions and withdrawal behaviours. These differences are represented by a model of emotional processing in which the frontal cortex plays a key role. Evidence supporting this model has been obtained from studies concerning asymmetry in prefrontal EEG alpha activity. Positive mood or reactions have been shown to be associated with relatively greater left prefrontal activity and negative mood or reactions with relatively greater right prefrontal activity. The results of recent neuroimaging studies suggest that negative affect typically elicits activation on the right prefrontal cortex, amygdala, and insula, and the left prefrontal cortex is associated with positive emotions [58]. The right prefrontal cortex may be critically involved in the response to stress, since it is a fundamental component of both the emotional and vigilance networks. Some studies suggest that high levels of right-sided prefrontal activation are associated with a negative affective style and weakened immune system. For example, Davidson [59][60] has reported that differences in prefrontal activity asymmetry reflect individual differences in affective styles. Also, the prefrontal cortex may mediate the extent to which psychosocial stress affects mental and physical health [61][62]. Differences appear to exist in how activity of the left and right cortical hemispheres affects ANS functioning. Moreover, the extent of this asymmetry has been suggested to vary under conditions of chronic stress [63]. Similar findings are reported for stress related emotions, with preferential right hemispheric activation in the frontopolar region that can be associated with electrodermal activity in anxious subjects [64]. The asymmetric analysis of the frequency-band powers in the EEG measured at the prefrontal cortex has been generally applied in previous stress studies. Particularly, the alpha band asymmetry is commonly used as indicator of the personal level of stress. Alpha waves reflect a calm, open, and balanced psychological state with a decrease in alpha wave activity during stress. Alpha wave training attempts to alleviate stress by inducing a state of relaxation. This involves removing or reducing habitual tendencies to respond to stressful situations with tension and anxiety [65].

## 2.2. Experimental setup

A deepened investigation was conducted to identify a set of devices enabling the measurement of the psycho-physical parameters mentioned above. We analysed different devices utilized in studies similar to the present. Besides technical considerations (e.g. type and quality of sensed signals, available software packages), the analysis was driven by other qualitative criteria: intrusiveness, reliability, and the ease of use were key features taken into account in the selection of the devices.

Non-contact measurement of the HR and the HRV could greatly simplify data acquisition, making such measurements easily available in non-clinical scenarios (e.g., driver monitoring [66], man-machine interaction monitoring [67]). Furthermore, non-contact systems could allow effective remote monitoring of patients. Several methods have been studied with this aim: the HR from speech [68], thermal imaging [69], microwave Doppler effect [70][71], and imaging photo plethysmography [66][72][73][74]. Imaging photo plethysmography (iPPG) is based on similar principles as the detection of finger pulse amplitudes. The heartbeat initiates the pulse wave and it travels through the arterial vascular system reaching various parts of the body. Here, the pulse wave determines a short-termed change of blood volume in the observed skin region: the intensity of the absorbed light depends on this volume, likewise the finger pulse measurement (standard plethysmography) [75]. It is therefore not surprising that several studies have been performed to evaluate whether the heart rate can be assessed from video streams [76], thus avoiding the use of wearable sensors. Concerning the signal acquisition in iPPG, video sequences are usually taken from a subject's face due to the high blood supply and the imaging simplicity. Both webcams [73][77][78][79][80] and conventional video cameras [81][82][83][84] have been used for this task. Webcams provide low-cost and easily available setups, whereas standard cameras are expected to produce better quality signals with higher spatial and temporal resolution, along with extended spectral capabilities. Tayibnapis et al. [66] utilized an infra-red camera in order to capture the driver's facial images. Zhao et al. [85] utilized a near-IR camera in order to extract the HR and the respiratory rate in both day and night light conditions. Authors also conclude that the HR and the RR can be extracted using single channel images. Typically, the strongest plethysmography signal is contained in the green channel [86]. This result is due to the fact that (oxy-)haemoglobin absorbs green light better than that red. Moreover, green light penetrates deeper into the skin than the blue light [82]. It stands to reason that using spectrally tuned band can significantly enhance iPPG signal.

Current methods to collect respiration data include the use of respiration belts, measurement of impedance through ECG electrodes, spirometers, or visual observation/counting. These techniques have drawbacks that limit the frequency and convenience of the respiratory monitoring. The large diffusion of wearable devices has stimulated interest in monitoring athlete training, with the aim of maximizing performance,

and minimizing the risk of injury and illness [87]. In these field, chest belts are very common choice and we resorted to adopt them. Measurement of the RR occurs through an embedded capacitive sensor composed of layers of conductive fabric, foam, and flexible mylar. Based upon the principle of a strain gauge sensor, thoracic expansion and contraction cause size differentials that induce changes in capacitance because of resultant changes in impedance. The change in impedance is manifested as a change in waveform signal amplitude represented as a sine wave with downward and upward deflections indicating chest expansion (increased impedance) and contraction (decreased impedance), respectively.

GSR can be measured by different methods. In general, GSR sensor measures the real-time skin conductance which is related to the sweat gland activity depending on emotional response and environmental condition [88][89]. GSR is typically acquired in hand fingers.

The majority of clinical studies use the EEG channels from hair-bearing scalp areas. However, this method requires the use of a conductive gel and an appropriate preparation procedure, which are particularly inconvenient for users. Indeed, the EEG recordings from hairless regions such as the forehead, or behind or inside the ear, would be more suitable for long-term monitoring in daily life. For this reason, there simple headband with sensors positioned in the hairless region are of common usage for various applications in well-being and fitness field.

On the basis of what we have just reported, the following set of the non-medical commercial devices were used:

1. Gigabit Ethernet camera with a CMOS monochrome sensor (UI-5240SE-NIR-GL, IDS GmbH DE). The camera was operated at 133 fps with an image matrix of 352 x 224 pixel, 8 bits/pixel. In order to enhance the plethysmographic signal, the camera mounted an optical band-pass filter cantered at 560 nm with a bandwidth of 40 nm.
2. BioHarness 3 Zephyr chest belt (Medtronic Inc, USA) for the respiratory rate measure with a frequency rate of 1 Hz.
3. Shimmer3 GSR (Shimmer, Ireland) for the galvanic skin response signal acquisition with a frequency rate of 256 Hz.
4. Muse 2 Headband (InteraXon Inc, USA) for the EEG signal acquisition.

A medical device was also used as ground truth for the ECG signal. To this end we adopted the EG05000 Medlab ECG Five Channel Module (Medlab GmbH, DE) operating at 300 Hz that it was employed to acquire the ECG lead sampled at 300 Hz.

All these sensors were connected to an Apple Mac Minicomputer (Intel Core i7 dual-core processor clocked at 3 GHz, 16 GB RAM, and 500 GB SSD).



Though each sensor had its own software tools (apps and libraries), the management of the integrated platform made it necessary to design ad hoc instruments so as to record signals simultaneously from the different source and store them with proper time stamping. This occurred by ad hoc wrappers written in C /C++ languages.

MATLAB (Mathworks, USA) was extensively used as the data analysis platform for all signals collected during the experimental sessions. Several MATLAB toolboxes were exploited, such as the Statistics and Machine Learning toolbox or the Curve Fitting toolbox.

During the acquisition procedures (Figure 4), subjects were sitting still in front of the computer monitor at a distance of about one meter. The chair had a headrest to contain head motion and make the recording comfortable for the volunteer. The subject face was illuminated by a white LED light source.

Subsequently, all the devices were positioned and calibrated. The ECG lead I had four electrodes corresponding to left and right arm, left and right limb. The camera was positioned on a tripod at a distance of about one meter from the subject. The BioHarness 3 Zephyr chest belt was positioned around the chest with the centre line of the device directly under the armpit. The two electrodes of the Shimmer3 GSR were positioned on the palmar surface of the proximal phalanx of the index and of the middle finger, respectively. The Muse 2 Headband is positioned as a normal headband.

After the sensor's calibration was asked to the subject to close his/her eyes and to relax. The subject's signals were recorded for five minutes in resting state. Subsequently, signals were acquired during a mild cognitive activation induced by the performed of a test.

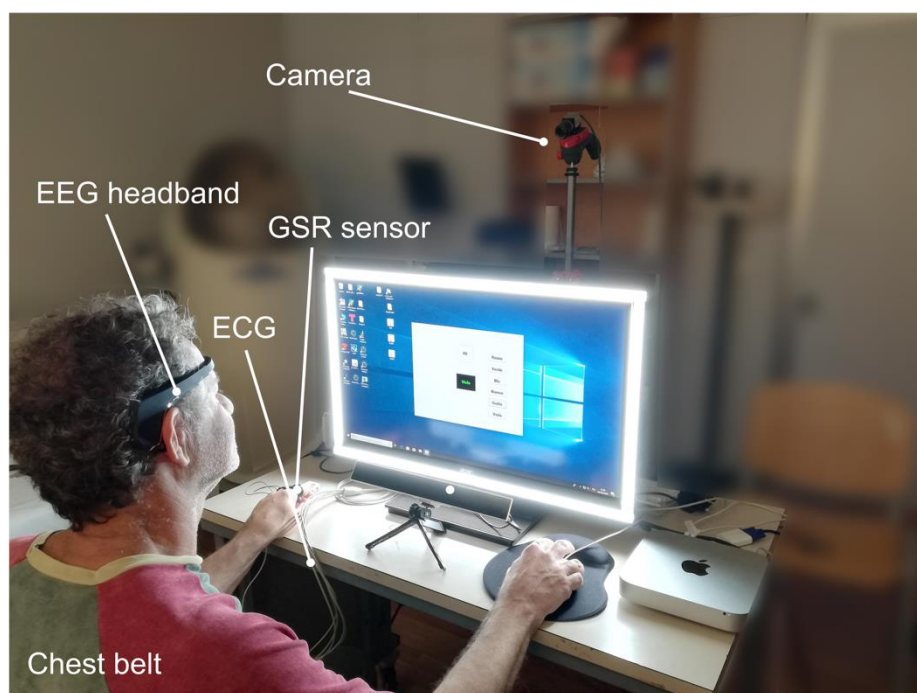


Figure 4 Experimental setup.

## 2.3. Testing protocol

Different laboratory procedures can be used to reliably induce stress in human research participants [90][91][92]. The most popular methods includes: the Stroop colour word test (SCWT), that will be described later [93][94][95][96][97][98], the Trier social stress test where subjects perform count backwards [90][99][100][101], and the Montreal imaging stress task where subjects perform basic math operation [102][103].

Taking into account that we are mainly focusing in moderate (non-maximal) personalized stimulations, we resorted to adopt the Stroop colour word test.

The SCWT is a neuropsychological test extensively used for both experimental and clinical purposes. It assesses the ability to inhibit cognitive interference, which occurs when the processing of a stimulus feature affects the simultaneous processing of another attribute of the same stimulus [93]. In the most common version of the SCWT, which was originally proposed by Stroop in the 1935, subjects are required to read three different tables as fast as possible. Two of them represent the “congruous condition” in which participants are required to read names of colours (henceforth referred to as colour words) printed in black ink (W) and name different colour patches (C). Conversely, in the third table, named colour-word (CW) condition, colour-words are printed in an inconsistent colour ink (e.g., the word “red” is printed in green ink). Thus, in this incongruent condition, participants are required to name the colour of the ink instead of reading the word. In other words, the participants are required to perform a less automated task (i.e., naming ink colour) while inhibiting the interference arising from a more automated task (i.e., reading the word) [94][95]. This difficulty in inhibiting the more automated process is called the Stroop effect [93].

In this project a portable version of the SCWT was implemented by an ad hoc Java app (Figure 5). We designed two different versions of the test:

1. Test A - a congruent version of the SCWT, where the font colour always matched the displayed colour name. The time limit to answer each question was set to two seconds and the overall test lasted two minutes.
2. Test B - an incongruent version of the SCWT, where the font colour not matched the displayed colour name. The time limit to answer each question was set to two seconds and the overall test lasted three minutes.

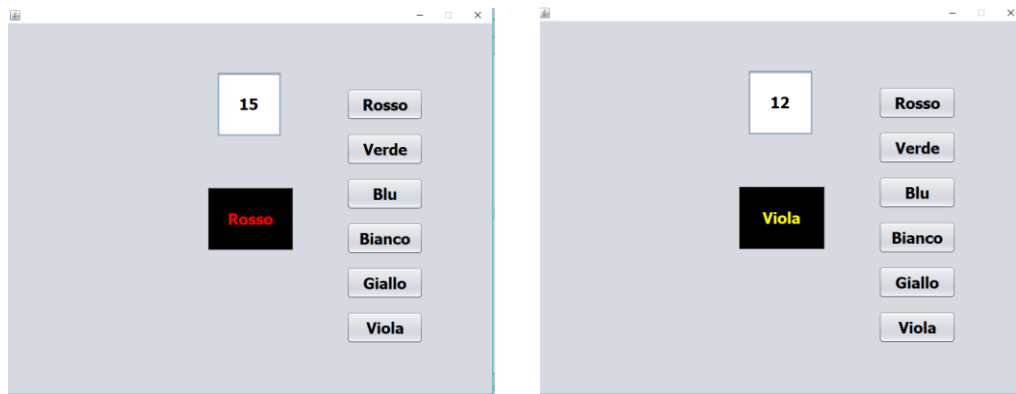


Figure 5 Test A on the left and Test B on the right.

## **Chapter 3. Heart rate and heart rate variability from single-channel video**

### **3.1. Introduction**

Several methods have been proposed to extract the blood volume pulse (BVP) signal from face movies. A widely adopted framework is based on blind source separation (BSS) techniques. The video signal is modelled as a mixture of contributes including BVP, motion artefacts, and external illumination changes. Poh et al. [73] processed the red-green-blue (RGB) video components by Independent Components Analysis (ICA) to enhance BVP. An alternative approach based on the BSS by Principal Components Analysis (PCA) was suggested by Lewandowska et al. [80]. The BSS was applied to RGB video components providing acceptable BVP estimation under the assumption of very small face motion and low image noise. To overcome such BSS (ICA- and PCA-based) limitations, several researchers investigated alternative processing methods. Wang et al. [24] exploits image redundancy to counteract the effect of the face movement. Feng et al. [79] adopted a simplified model of the optical properties of the skin to compensate for head motion. Tarassenko et al. [83] proposed an iPPG system exploiting autoregressive modelling of video time series to compute the HR together with the respiratory rate and the Oxygen saturation. Though video signal intensity is the most utilized source of information to detect BVPs, a different method based on head motion related to BVP propagation is reported in [104].

The assessment of the HRV from video is usually more demanding than measurement of the HR. In fact, the HR estimation only requires BVP detection so as to compute the average number of pulse per minute, the HRV assessment requires a precise temporal localization of pulses. In principle, most methods for the HR assessment from video can be adapted to estimate the HRV. Video sequences are usually processed to detect blood volume pulses and obtain a tachogram (i.e., the time series of inter-beat interval duration), similarly to what done in ECG-based analysis. In particular, the tachogram can be analysed both in time domain and/or in frequency domain [105]. For example, in [73] iPPG tachogram is used to compute standard HRV descriptors both in the time and the frequency domain. High correlation with parameters derived by standard photoplethysmography (PPG) on 15 subjects is reported. Another solution based on Zero-phase Component Analysis has been reported by Iozzia et al. [106] to evaluate the suitability of iPPG to assess autonomic response. Tayibnapis et al. [66] applied the PCA to estimate HR and HRV signals from an

RGB video. These signals were then used with a set of facial features to detect fatigue through a Support Vector Machine.

In general, the HR estimation from iPPG is in good agreement with reference techniques. In particular, it was shown that the HR estimated from the iPPG highly correlates (Pearson coefficient  $> 0.90$ ) with the HR extracted from standard ECG signals [73]. Standard PPG, due to its simplicity, is the most common reference in iPPG studies. Though PPG provides accurate HR measurements, the gold standard for HRV assessment remains ECG recording that allow a fine localization of the heart beat [105]. PPG seems a viable surrogate of ECG for healthy subject at rests, but its performances tends to worsen in exercise and in diseased people [107][108]. In addition, it must be pointed out that, up to now, experimental results are from small sized populations, data acquisition usually occurring in highly idealized conditions. In particular, short-term analysis is considered by most researchers with video recordings lasting non longer than 60 s, instead of the usual 5 minutes recommend in [105], usually capture with a subject staying still in front of a camera.

Obviously, when acquisition constraints are relaxed, several factors can alter the iPPG signal and degrade the performances of processing algorithms. For example, subject movements can be a source of troublesome artefacts. Rigid motions can be somehow compensated by proper tracking of the region of interest but the effects of non-rigid facial movements can be hardly removed. Rapid changes of environmental lighting can also interfere with iPPG and this need to be taken into account when setting up an iPPG system. Imaging parameters such as spatial and temporal sampling and the sensor spectral response impact on image quality. Though spatial resolution does not seem a major iPPG problem (mostly because iPPG signal is usually averaged on large facial areas), temporal resolution is expected to affect BVP localization with a possibly relevant impact on the HRV analysis. In this respect, it is worth remembering that in standard ECG-based HRV measurements sampling rates lower than 100 Hz are discouraged [105]. Though iPPG is based on different principles than ECG, temporal sampling requirements should be more deeply investigated.

Through this analysis we report on a novel iPPG methodology to monitor the HR and the HRV of normal subjects [109]. Based on the previous considerations, the work described in this chapter has a twofold aim: a) investigate the use of ICA pre-processing of spectrally-tuned single-channel video streams to enhance iPPG signal by exploiting multiple spatial measurements, b) analyse the impact of imaging parameters on the HR and the HRV descriptors estimation. In next sections, after describing the experimental setup, we report on the estimation of HR and HRV parameters (computed both in time and frequency domain) in varying experimental conditions. The iPPG-derived parameters are compared with corresponding descriptors derived from simultaneous ECG recording.

### 3.2. Blood volume pulse from video

The propagation of blood volume pulses makes skin reflectance changing in time, which, in turn, changes the colour of video recordings and this is the primary source of information we consider to detect HR from face videos. However, colour changes can be produced by other physiological processes, voluntary subject movement as well as ambient factors. In this work, we refer to the experimental setup sketched in Figure 6, where an individual is sitting in front of a camera with the face uniformly lit by a stable white LED lamp with no other (natural or artificial) light sources. On this ground, we will focus mainly on subject-related effects.

Among these, respiration is responsible of small cyclic head and neck movements that can be detected by imaging face regions including motion sensitive features (e.g., the borders of lips). In addition, involuntary movements (e.g., muscular tremor and eye blinking) should be taken into account as potential source of interference. Voluntary head motion that may be related to a wide variety of activities (e.g., speaking, tracking moving objects) can be expected to introduce relevant picture changes interfering with the recording of vital signs. In this work, the BVP signal is extracted from a set of observations of the iPPG video taken in different facial regions. Let assume that, for a fixed wavelength, a iPPG signal  $g(t)$  is obtained at time  $t$  by spatially averaging the image intensity in a region of interest (ROI). We assume that averaging makes camera noise negligible.

According to a widely accepted model for iPPG [110],  $g(t)$  is produced by specular and diffuse reflections of the incident light  $I_0(t)$ . Separating continuous and time varying, zero mean, contributes, one can write:

$$g(t) = I_0(1 + i(t))(k + \mu m(t) + \sigma p(t)) \quad (1)$$

where  $I_0(t) = I_0(1 + i(t))$ ,  $k$  accounts for all contributes to constant reflection,  $m(t)$  describes changes of specular reflection, and  $p(t)$  relates to changes of diffuse reflection. The coefficients  $\mu$  and  $\sigma$  quantify the change of skin reflection and the strength of BVP pulse, respectively. It is worth noting that the term  $p(t)$  is the only contribute pertaining BVP. Equation 1 can be simplified by assuming that the time varying components are small compared to continuous ones and that their cross products can be neglected:

$$g(t) \approx I_0 k + \mu I_0 k i(t) + I_0 \mu m(t) + I_0 \sigma p(t) \quad (2)$$

We can therefore conclude that the iPPG signal is, approximately, a linear combination of three contributes:  $i(t)$  related to illumination changes that, in our case, is related to subject

motion with respect to the light source<sup>1</sup>,  $m(t)$  that describes specular changes of skin surface due to subject motion, and  $p(t)$  that must be imputed to diffused reflection and accounts for BVP contribute. The signal  $i(t)$  is often estimated from RGB videos which provides three different combinations of the underlying.

In the following we will consider monochrome images acquired with a narrow-band filter centred on haemoglobin absorption peaks. Multiple iPPG signals are generated from different ROIs in regions having high vascularization. Assuming that  $i(t)$ ,  $m(t)$ , and  $p(t)$  are approximately uniform among the ROIs we rewrite Equation 2 for each region ROI:

$$g_j(t) = I_0 k_j + I_0 c_j i(t) + I_0 \mu_j m(t) + I_0 \sigma_j p(t) \quad (3)$$

where  $k_j$ ,  $\mu_j$ , and  $\sigma_j$  are the constant reflection, the specular and the diffuse reflection in the  $j$ -th ROI, respectively. Based on Equation 3 we exploit the spatial dependence of  $g(t)$  to enhance BVP estimation by means of ICA.

### 3.3. Materials and methods

#### 3.3.1. Study's participants

Thirty healthy participants (12 females and 18 males with mean age of 39.17 years, range from 22 to 61) were recruited for voluntary participation in this study. The presented examination was part of a larger study aiming at evaluating the possibility of using imaging techniques for individual self-assessment and self-monitoring of cardio-metabolic risk [109]. The experimental protocol, conducted in Pisa, Italy, was approved (September 10th, 2015) by the local Ethical Committee of Pisa, Italy (Study 213/2014) and received the Ethical Clearance certification (0086129, November 11th, 2014) by the Italian National Commission for Research Ethics and Bioethics. Written informed consent was obtained from all subjects included in this study. The study protocol is compliant with the European Union General Data Protection Regulation [111].

#### 3.3.2. Experimental setup and video acquisition

The subjects were sitting still in front of the camera at a distance of about one meter (Figure 6). The chair had a headrest to contain head motion and making the video recording comfortable for the volunteer. After a three minutes rest the subject's face was recorded for five minutes. The subjects were illuminated by a white LED light source. The acquisitions

---

<sup>1</sup> In principle, small cyclic head movements due to Newtonian reaction to blood influx may results in related subtle cyclic changes of skin reflectance. However, their impact on  $m(t)$  in a uniform skin patch, not including motion sensitive features, can be neglected.

were performed through the camera and the ECG system described in section 2.2. All the videos and the ECG recordings were acquired and analysed through a custom software written in C++.

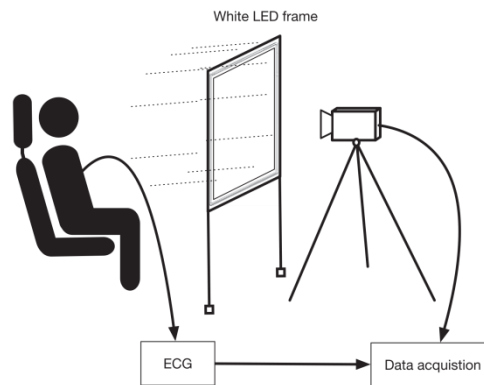


Figure 6 Data acquisition setup.

### 3.3.3. ECG signal analysis

The reference data were obtained by extracting the tachogram from the ECG signal. To this end, the Inter Beat Intervals were defined following QRS complex detection [112]. To remove possible artefacts, the interval beat intervals (IBIs) were processed by the NC-VT algorithm [113] with a tolerance of 30 %, all the IBIs with a duration less than 200 ms being removed. The signals were analysed both in the time domain and in the frequency domain. Concerning the time domain, we calculated the average time between adjacent normal heartbeats (NN) and its standard deviation (SDNN). Concerning the frequency domain, analysis of the HRV was performed by the power spectrum density (PSD) estimated by the Lomb-Scargle periodogram [114]. This method is able to cope with non-uniformly sampled data even in presence of large gaps, which makes tachogram interpolation unnecessary. According to the standard definition of the HRV frequency bands [105], the low frequency (LF) and the high frequency (HF) were calculated as the area under the PSD curve corresponding from 0.04 Hz to 0.15 Hz and from 0.15 Hz to 0.4 Hz, respectively [113]. The LF component reflects both sympathetic and parasympathetic actions, the HF component reflects parasympathetic action, and the LF/HF ratio is a measure of the sympatho/vagal balance [113]. The features extracted are the most commonly used in the literature for short-term analysis [115][116].



### 3.3.4. ROI's selection

For each video sequence, the Cascading Classifiers algorithm (OpenCV v. 3.0.0) was used to detect the face of the subject and locating the centre of each eye in the first frame of the recorded video, the related distance being named  $d$ . Three different regions of interest (ROIs) were selected (Figure 7). ROI definition was driven by anatomical and empirical considerations. In particular, ROIs were placed in highly vascularized regions of the face. We also tried to have a large integration area so as to reduce the impact of image noise on BVP estimation. In addition, ROI sizes were constrained to reduce interference with regions interested by marked involuntary movement, such as eyes and lips. The first ROI (ROI1) was a rectangle of the forehead area and the others (ROI2 and ROI3) were squares on the right and left cheek, respectively. The ROI on the forehead was  $d \times 3/8 d$  placed at  $5/8 d$  above the inter-pupillary line. The two ROIs on the cheeks had a fixed side  $l = 20$  pixel (about 20 mm on the subject face) placed at  $d/2$  under the pupil.

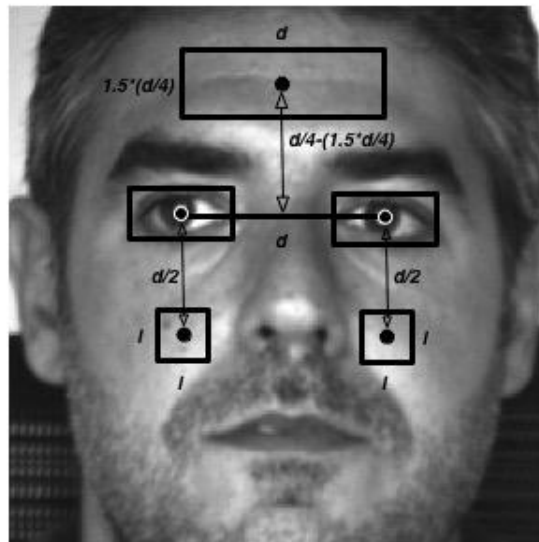


Figure 7 Dimensions and positions of the ROIs (ROI1 on the forehead area, ROI2 on the right cheek, and ROI3 on the left cheek).  $d$  is the distance between the centres of the rectangles around the eyes and  $l$  is the side of the square that is set to 20 pixel.

### 3.3.5. Video signal analysis

In order to extract the HR and the HRV parameters from the video, for each frame, grey levels were averaged in each ROI, which provided three iPPG signals ( $v_1(t)$ ,  $v_2(t)$ , and  $v_3(t)$ ). Then, two procedures were implemented to recover the pulse signal:

1. M1 - Each  $v_i(t)$  was filtered using a FIR band-pass filter implemented via Hamming window, with lower cut-off at 0.75 Hz (45 bpm) and upper cut-off at 2 Hz (120 bpm). Such a band is well suited for normal subjects at rest or in during moderate activity,

the bandwidth can be adapted during intense physical activity or in presence of disease. The signals  $x_1(t)$ ,  $x_2(t)$ , and  $x_3(t)$  were so obtained.

2. M2 - As illustrated in Figure 8, each  $v_i(t)$  was detrended by subtracting a 2 s time average and the signals  $y_1(t)$ ,  $y_2(t)$ , and  $y_3(t)$  were obtained. Following whitening, the three detrended signals were jointly processed through Independent Component Analysis as implemented in FastICA algorithm [117]; three new sequences  $y_A(t)$ ,  $y_B(t)$ , and  $y_C(t)$  were so obtained, each representing a different signal contribution. The spectra of  $y_A(t)$ ,  $y_B(t)$ , and  $y_C(t)$  were computed via Fast Fourier Transform. The BVP-related component was defined as the one with the highest peak in the range from 0.75 Hz to 2.0 Hz and was then filtered as in M1 and  $y_S(t)$  was so obtained. It is worth noting that while the algorithm used (that is the FastICA) is not new it was applied to three portions of a single-channel video, in opposition to what is usually done (one portion of a three-channel video).

In both procedures, after filtering, the BVP peaks were detected using a multi-scale algorithm [118] and local quadratic interpolation. Finally, the tachogram for each BVP peak sequence was computed. As described for ECG analysis, possible artefacts were removed by the NC-VT algorithm. Similarly to the ECG signal, the NN and the SDNN values were extracted from the tachogram. The frequency domain analysis was performed by PSD estimation by Lomb-Scargle periodogram and the LF, HF, and LF/HF values were calculated from the PSD.

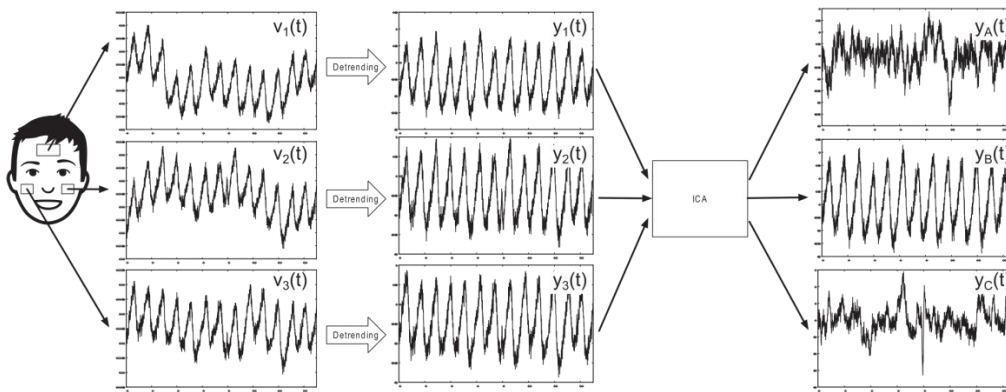


Figure 8 Procedure based on ICA.

### 3.3.6. Recording length

The videos with different duration were compared with corresponding ECG portions. In particular, pieces spanning one, two, three, four, and five minutes (entire recording) were extracted from the available sequence.

### 3.3.7. Video frame rate

The original videos were recorded at 133 fps. To test the effect of lower sampling rates as those often reported in literature, each video was down-sampled from the native 133 fps to 66 fps and 33 fps. The ROIs signals computed from down-sampled videos were interpolated with a cubic spline function and resampled at 133 Hz.

## 3.4. Results

The data analysis was based on absolute errors observed by comparing iPPG with ECG. The HR and HRV descriptors estimated from video (both by M1 using band-pass filter only and M2 based on ICA pre-processing) were compared to corresponding parameters obtained from ECG. The mean, standard deviation, minimum, and maximum values of the data obtained by the ECG and the video signal are reported in next sections for the various descriptors. For each of them, we give the mean absolute error (MAE), with standard deviation ( $SD_{AE}$ ), minimum ( $AE_{min}$ ), and maximum ( $AE_{max}$ ). The median of signed errors ( $M_e$ ) is also reported as a robust indicator of systematic error. In addition, as correlation coefficients are commonly used by many researchers, Pearson coefficient  $r_p$  and Spearman's rho  $\rho_s$  are also reported for completeness.

### 3.4.1. Analysing video at the native frame rate

The mean, standard deviation, minimum, and maximum values of the data obtained by the ECG and the video signal analysis (both obtained from M1 and M2) are reported in Table 2. The observed MAEs and correlations obtained from each ROI (ROI1, ROI2, and ROI3) are reported in Table 3. With respect to NN, both the video signals and the ICA perform well against the ECG showing low values of the MAEs, with the smallest ones provided by  $x_1(t)$  (3.812 ms) and  $y_s(t)$  (3.822 ms), and strong correlations ( $x_1(t)$ ,  $x_2(t)$ ,  $x_3(t)$ , and  $y_s(t)$ ) with  $r_p \geq 0.999$  and  $\rho_s \geq 0.995$ . Similarly, for SDNN, LF, HF, and LF/HF, the smallest MAEs values and strongest correlations are observed in  $x_1(t)$  and  $y_s(t)$ ; instead, the video signals from the cheeks ROIs ( $x_2(t)$  and  $x_3(t)$ ) are more prone to artefacts, so that the extracted parameters exhibits a weaker correlation with the ECG data than signal from the forehead ROI. For  $x_1(t)$  the MAEs observed values are 5.643 ms for SDNN, 3.303 % for LF, 6.556 % for HF, and 0.657 for LF/HF; for  $y_s(t)$  we have 5.853 ms for SDNN, 2.775 % for LF, 4.372 % for HF, and 0.482 for LF/HF. For  $x_1(t)$  the observed correlations are  $r_p = 0.960$  ( $\rho_s = 0.903$ ) for SDNN,  $r_p = 0.878$  ( $\rho_s = 0.813$ ) for LF,  $r_p = 0.786$  ( $\rho_s = 0.779$ ) for HF, and  $r_p = 0.842$  ( $\rho_s = 0.779$ ) for LF/HF. For  $y_s(t)$  the observed correlations are  $r_p = 0.961$  ( $\rho_s = 0.928$ ) for SDNN,  $r_p = 0.916$  ( $\rho_s = 0.895$ ) for LF,  $r_p = 0.885$  ( $\rho_s = 0.884$ ) for HF, and  $r_p = 0.931$  ( $\rho_s = 0.876$ ) for LF/HF. All the parameters obtained from  $y_s(t)$  (M2) were strongly correlated with the reference ones, the correlation being larger than those observed for M1 in all cases. Scatter plots of parameters computed

from  $x_1(t)$  and  $y_S(t)$  and the corresponding ECG ones are summarized in Figure 9 (time domain) and Figure 10 (frequency domain).

Table 2 Mean, standard deviation, minimum, and maximum values of the data obtained by ECG and video signal analysis at the native frame rate.

Parameter	Metrics	ECG	$x_1(t)$ ROI1	$x_2(t)$ ROI2	$x_3(t)$ ROI3	$y_S(t)$ ICA
NN (ms)	Mean	891.281	894.241	890.732	893.661	894.521
	SD	141.412	142.372	143.170	142.012	142.575
	Min	666.470	664.189	648.958	669.968	664.090
	Max	1127.100	1133.250	1133.390	1127.130	1132.080
SDNN (ms)	Mean	55.663	58.885	71.285	69.498	59.667
	SD	20.302	19.902	23.646	19.637	19.684
	Min	26.189	32.928	41.714	37.094	30.944
	Max	113.834	122.372	123.666	124.734	117.419
LF (%)	Mean	29.153	29.112	30.142	30.039	29.404
	SD	9.035	8.901	8.309	8.008	8.730
	Min	11.780	13.894	12.267	12.626	14.068
	Max	43.315	41.532	48.140	45.490	41.574
HF (%)	Mean	25.404	26.911	33.091	32.916	27.692
	SD	12.742	12.899	12.711	12.755	12.286
	Min	4.042	9.182	13.560	11.980	8.447
	Max	56.621	61.521	67.840	68.999	58.625
LF/HF	Mean	1.752	1.409	1.099	1.114	1.377
	SD	1.814	0.930	0.609	0.672	0.970
	Min	0.275	0.235	0.181	0.183	0.259
	Max	9.767	4.427	2.476	3.349	4.910

Table 3 Absolute errors (mean, standard deviation, minimum, and maximum), median of signer errors, Pearson coefficient, and spearman's obtained using M1 and M2, respectively.

Parameter	Metrics	$x_1(t)$ ROI1	$x_2(t)$ ROI2	$x_3(t)$ ROI3	$y_S(t)$ ICA
NN	MAE (ms)	3.812	6.124	4.309	3.822
	SD <sub>AE</sub> (ms)	4.174	6.253	4.441	4.076
	AE <sub>min</sub> (ms)	0.332	0.075	0.114	1.210
	AE <sub>max</sub> (ms)	24.311	24.217	23.912	24.424
	M <sub>e</sub> (ms)	2.193	1.512	2.156	3.157
	$r_p$	0.999	0.998	0.999	1.000
	$\rho_s$	0.998	0.995	0.996	0.998
SDNN	MAE (ms)	5.643	16.653	14.750	5.853
	SD <sub>AE</sub> (ms)	3.197	11.851	11.722	4.076
	AE <sub>min</sub> (ms)	0.028	0.251	1.593	1.500
	AE <sub>max</sub> (ms)	10.604	53.427	47.773	17.355
	M <sub>e</sub> (ms)	4.202	15.164	12.231	3.867
	$r_p$	0.960	0.830	0.794	0.961
	$\rho_s$	0.903	0.757	0.723	0.928
LF	MAE (%)	3.303	4.666	5.459	2.775
	SD <sub>AE</sub> (%)	2.810	4.417	5.421	2.341
	AE <sub>min</sub> (%)	0.068	0.181	0.110	0.262
	AE <sub>max</sub> (%)	11.261	15.715	21.345	8.361
	M <sub>e</sub> (%)	0.783	0.205	0.579	0.758
	$r_p$	0.878	0.730	0.596	0.916
	$\rho_s$	0.813	0.685	0.637	0.895
HF	MAE (%)	6.556	11.716	9.605	4.372
	SD <sub>AE</sub> (%)	5.324	8.460	8.978	4.684
	AE <sub>min</sub> (%)	0.353	0.076	0.359	0.592
	AE <sub>max</sub> (%)	20.238	36.965	33.823	18.450
	M <sub>e</sub> (%)	3.324	8.046	5.270	1.436
	$r_p$	0.786	0.529	0.638	0.885
	$\rho_s$	0.779	0.519	0.643	0.884
LF/HF	MAE	0.657	0.925	0.806	0.482
	SD <sub>AE</sub>	0.994	1.444	1.213	0.927
	AE <sub>min</sub>	0.007	0.003	0.001	0.002
	AE <sub>max</sub>	5.339	7.978	6.418	4.857
	M <sub>e</sub>	-0.187	-0.356	-0.242	-0.046
	$r_p$	0.842	0.512	0.827	0.931
	$\rho_s$	0.779	0.487	0.584	0.876

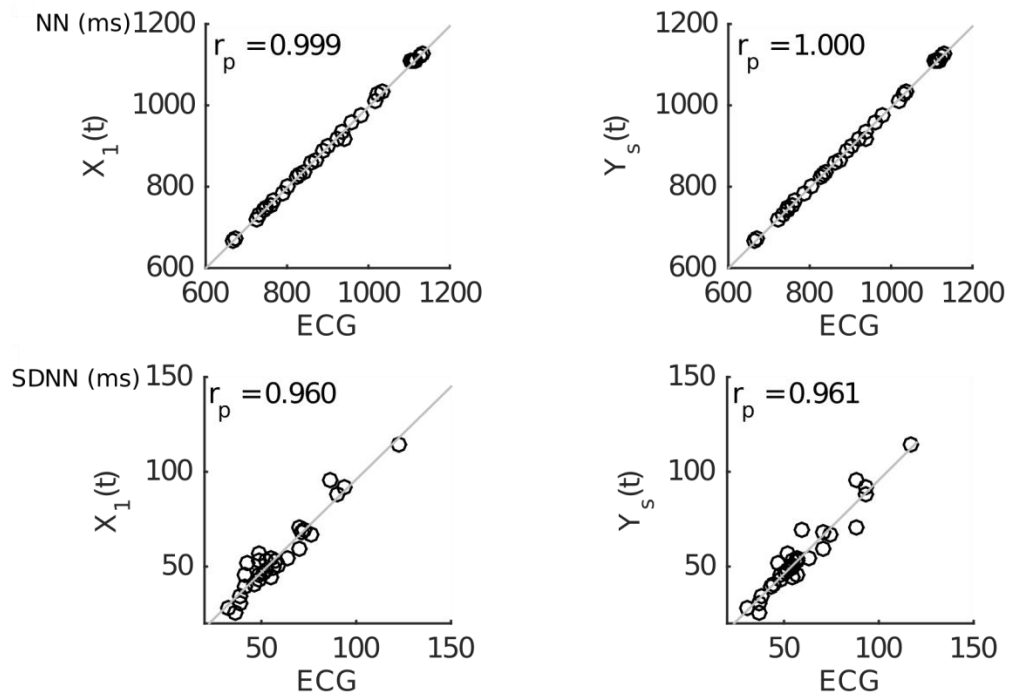


Figure 9 Scatter plots comparing NN and SDNN between the video signal and ECG signal. The video signal data are obtained from  $x_1(t)$  applying M1, on the left side, and from  $y_s(t)$  applying M2, on the right side.

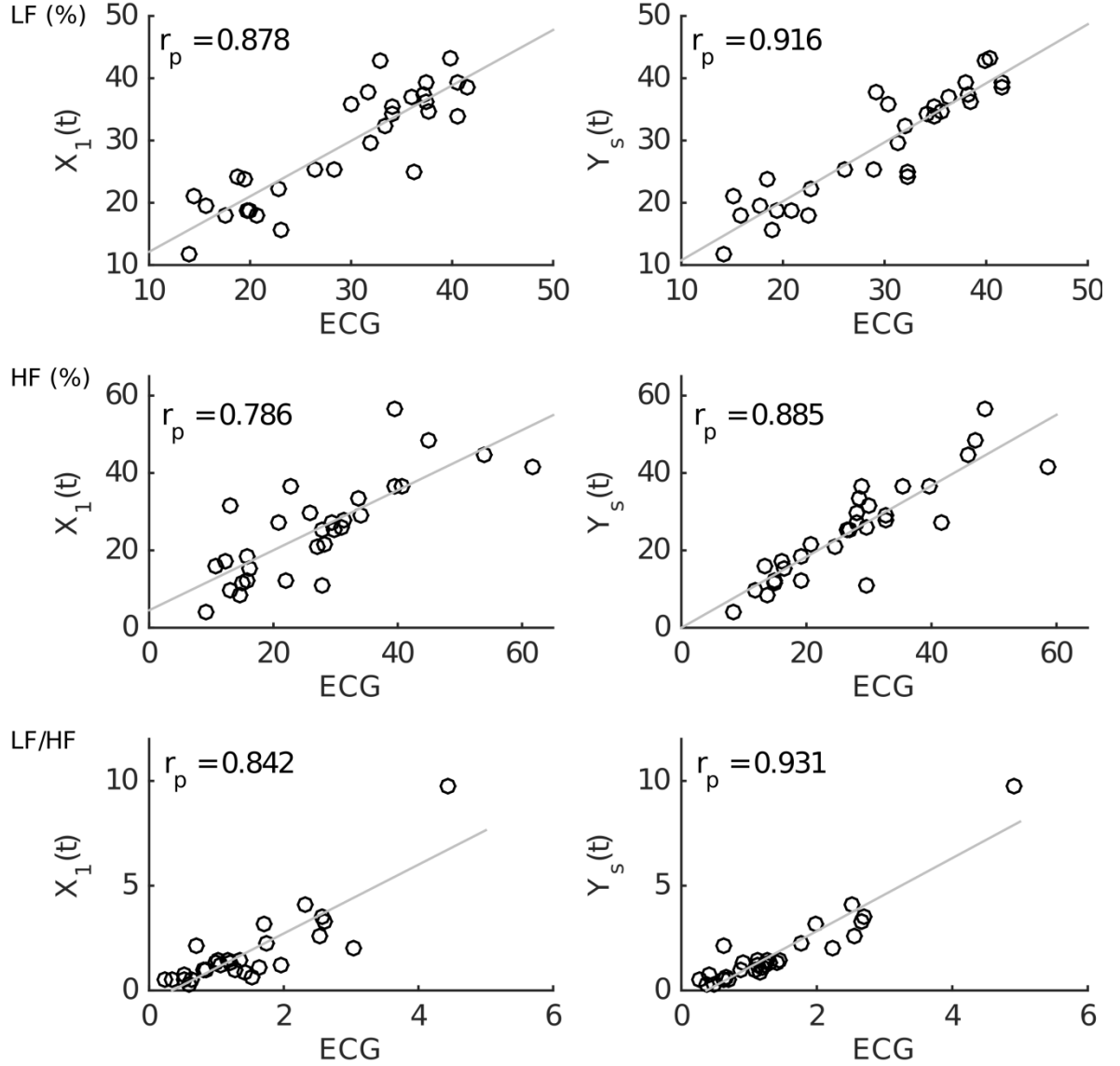


Figure 10 Scatter plots comparing LF, HF, and LF/HF values between the video signal and ECG signal. The video signal data are obtained from  $x_1(t)$ .

### 3.4.2. Varying the frame rates

In Table 4 and Table 5 we analysed the parameters estimated from the down-sampled video signals. In particular, we analysed the data obtained by M1 from the forehead signal  $x_1(t)$  (the best performing at 133 fps) and those obtained from  $y_s(t)$  by M2. The MAE values related to NN extracted by the down-sampled video signals for both M1 and the M2 do not exhibit appreciable differences (being all slightly larger than 3 ms) at different frame rates. High correlation coefficients are observed in all cases ( $r_p \geq 0.999$  and  $\rho_s \geq 0.998$ ). The SDNN behaviour shows increasing MAEs when frame rate decreases for both M1 and M2; in parallel, the correlation coefficients decrease at the slowest frame rate. Also in the frequency domain, the MAE values increase as frame rate decrease, this being more pronounced for HF and LF/HF. Generally, the MAE values and correlation coefficients

obtained for M2 are better than those obtained for M1. Furthermore, in the frequency domain, the contribution of the ICA is outstanding when the sampling rate is less than 100 fps.

Table 4 Mean, standard deviation, minimum, and maximum values of the data obtained by video signal analysis varying the frame rates.

Parameter	Metrics	$x_1(t)$ (ROI1) video rate (fps)			$y_S(t)$ (ICA) video rate (fps)		
		133	66	33	133	66	33
NN (ms)	Mean	894.241	892.987	892.381	894.521	893.122	893.041
	SD	142.372	142.134	142.009	142.575	142.130	142.373
	Min	664.189	662.083	661.517	664.090	661.946	661.843
	Max	1133.250	1131.990	1131.960	1132.080	1130.350	1130.680
SDNN (ms)	Mean	58.885	59.891	62.022	59.667	61.780	63.273
	SD	19.902	19.867	20.663	19.684	20.638	20.001
	Min	32.928	32.952	34.147	30.944	31.485	32.262
	Max	122.372	122.391	123.756	117.419	117.636	116.960
LF (%)	Mean	29.112	29.111	29.394	29.404	29.448	28.875
	SD	8.901	8.637	8.245	8.730	8.202	8.154
	Min	13.894	14.090	14.271	14.068	14.897	14.440
	Max	41.532	42.007	41.195	41.574	41.300	41.185
HF (%)	Mean	26.911	27.280	28.275	27.692	28.368	29.335
	SD	12.899	13.142	11.962	12.286	11.977	12.055
	Min	9.182	9.930	11.363	8.447	8.948	10.637
	Max	61.521	60.534	60.194	58.625	57.935	57.330
LF/HF	Mean	1.409	1.386	1.302	1.377	1.321	1.240
	SD	0.930	0.909	0.824	0.970	0.898	0.807
	Min	0.235	0.240	0.245	0.259	0.261	0.268
	Max	4.427	4.025	3.528	4.910	4.616	3.872



Table 5 Absolute errors (mean, standard deviation, minimum, and maximum, median of signed errors, Pearson coefficient, and spearman's obtained at three different video rates for  $x_1(t)$  and  $y_s(t)$  using M1 and M2, respectively.

Parameter	Metrics	$x_1(t)$ (ROI1) video rate (fps)			$y_s(t)$ (ICA) video rate (fps)		
		133	66	33	133	66	33
NN	MAE (ms)	3.812	3.337	3.526	3.822	3.136	3.251
	SD <sub>AE</sub> (ms)	4.104	3.993	4.263	4.008	3.840	3.929
	AE <sub>min</sub> (ms)	0.332	0.284	0.248	1.210	0.637	0.057
	AE <sub>max</sub> (ms)	24.311	22.713	22.643	24.424	22.872	22.512
	M <sub>e</sub> (ms)	2.193	1.991	1.812	3.157	2.227	1.868
	$r_p$	0.999	0.999	0.999	1.000	0.999	0.999
	$\rho_s$	0.998	0.998	0.998	0.998	0.998	0.998
SDNN	MAE (ms)	5.643	6.518	8.668	5.835	7.118	8.497
	SD <sub>AE</sub> (ms)	3.143	3.212	4.8273	3.502	3.810	4.732
	AE <sub>min</sub> (ms)	0.028	0.300	1.210	1.500	3.036	1.939
	AE <sub>max</sub> (ms)	10.604	12.596	23.027	17.355	19.968	20.645
	M <sub>e</sub> (ms)	4.202	5.500	8.137	3.867	6.137	7.786
	$r_p$	0.960	0.955	0.929	0.961	0.966	0.953
	$\rho_s$	0.903	0.887	0.874	0.928	0.937	0.911
LF	MAE (%)	3.303	3.179	3.365	2.775	3.311	3.094
	SD <sub>AE</sub> (%)	2.850	3.278	4.827	2.302	2.293	4.732
	AE <sub>min</sub> (%)	0.068	0.015	0.002	0.262	0.190	0.077
	AE <sub>max</sub> (%)	11.261	12.971	13.862	8.361	7.996	9.861
	M <sub>e</sub> (%)	0.783	0.629	0.525	0.758	0.774	0.534
	$r_p$	0.878	0.863	0.855	0.916	0.892	0.889
	$\rho_s$	0.813	0.827	0.828	0.895	0.872	0.863
HF	MAE (%)	6.556	6.580	7.083	4.372	4.851	5.571
	SD <sub>AE</sub> (%)	5.234	5.571	5.423	4.606	4.906	4.898
	AE <sub>min</sub> (%)	0.353	0.454	1.074	0.592	0.093	0.092
	AE <sub>max</sub> (%)	20.238	19.251	21.200	18.450	21.294	22.976
	M <sub>e</sub> (%)	3.324	3.157	4.304	1.436	2.937	3.944
	$r_p$	0.786	0.782	0.760	0.885	0.870	0.868
	$\rho_s$	0.779	0.766	0.774	0.884	0.871	0.850
LF/HF	MAE	0.657	0.657	0.708	0.482	0.539	0.620
	SD <sub>AE</sub>	0.978	1.037	1.126	0.912	0.975	1.084
	AE <sub>min</sub>	0.007	0.074	0.049	0.002	0.004	0.003
	AE <sub>max</sub>	5.339	5.742	6.239	4.857	5.151	5.895
	M <sub>e</sub>	-0.187	-0.196	-0.229	-0.047	-0.086	-0.197
	$r_p$	0.842	0.817	0.786	0.931	0.922	0.887
	$\rho_s$	0.782	0.739	0.739	0.876	0.844	0.828

### 3.4.3. Varying the acquisition time

We analysed the parameters obtained from the video signals (both M1 and M2) having different durations. As in the case of section 3.4.2, we considered  $x_1(t)$  and  $y_s(t)$  only. In Table 6 we reported the basic statistics for parameters measured from ECG,  $x_1(t)$ , and  $y_s(t)$ , respectively. In Table 7 the MAEs and the correlation coefficients are given. Both M1 and M2 show similar error pattern for NN: the MAE increases when the recording time decreases, difference being slightly more evident for durations less than or equal to 2 minute. High correlation coefficients were observed at any duration for  $x_1(t)$  ( $r_p \geq 0.998$  and  $\rho_s \geq 0.995$ ) and  $y_s(t)$  ( $r_p \geq 0.999$  and  $\rho_s \geq 0.994$ ). The MAEs of remaining parameters tend

to increase as the recording duration decreases, the HF showing the more pronounced dependence. As shown in Table 7, dependence on acquisition time was slightly more pronounced for the frequency domain descriptors. In any case, only for the shortest observation window (1 minute) we observed a MAE approximately doubled with respect to other durations.

Table 6 Mean, standard deviation, minimum, and maximum values of the data obtained by ECG and video signal analysis varying the acquisition time from M1.

Parameter	Metrics	5 min			4 min			3 min			2 min			1 min		
		ECG	$x_1(t)$ ROI	$y_S(t)$ ICA	ECG	$x_1(t)$ ROI	$y_S(t)$ ICA	ECG	$x_1(t)$ ROI	$y_S(t)$ ICA	ECG	$x_1(t)$ ROI	$y_S(t)$ ICA	ECG	$x_1(t)$ ROI	$y_S(t)$ ICA
NN (ms)	Mean	891.281	894.241	894.521	893.874	896.17	896.454	899.108	901.651	902.224	903.163	905.632	906.013	912.381	913.224	913.700
	SD	141.412	142.372	142.373	144.211	144.084	144.471	144.510	144.735	145.262	148.181	148.043	148.426	148.816	149.370	149.751
	Min	666.470	664.189	664.090	664.637	661.864	661.906	664.689	660.676	660.504	655.403	648.219	647.808	669.886	655.705	656.141
	Max	1127.100	1133.250	1132.080	1141.610	1139.580	1138.990	1149.250	1147.850	1147.000	1175.780	1167.650	1168.820	1183.200	1176.390	1176.090
SDNN (ms)	Mean	55.663	58.885	59.667	55.346	58.700	59.277	54.915	58.361	57.903	53.622	56.799	55.790	51.433	55.447	53.842
	SD	20.302	19.902	19.684	21.705	20.560	20.281	23.200	21.709	21.244	20.471	19.372	19.067	18.416	19.115	17.760
	Min	26.189	32.928	30.944	23.985	33.88	31.365	24.901	34.952	31.948	25.845	30.653	28.958	24.750	30.757	29.153
	Max	113.834	122.372	117.419	118.823	127.048	123.851	126.919	133.067	131.597	99.561	107.222	105.608	99.203	103.733	99.853
LF (%)	Mean	29.153	29.112	29.404	26.318	28.559	28.627	27.524	30.209	30.190	28.801	31.144	31.240	30.809	35.339	35.576
	SD	9.035	8.901	8.730	8.703	8.632	8.935	10.566	10.846	11.182	10.809	10.413	10.210	13.886	15.492	15.552
	Min	11.78	13.894	14.068	11.560	14.021	13.932	12.515	13.954	14.527	11.702	13.170	13.739	14.347	12.143	14.838
	Max	43.315	41.532	41.574	42.623	46.050	45.967	49.951	58.773	58.833	55.247	57.848	57.797	64.490	77.275	75.240
HF (%)	Mean	25.404	26.911	27.692	32.142	27.467	27.943	34.318	28.936	29.006	38.412	33.405	32.786	43.135	35.284	34.604
	SD	12.742	12.899	12.286	12.499	12.406	11.951	11.968	12.411	11.255	16.414	16.736	15.651	16.441	16.475	15.041
	Min	4.042	9.182	8.447	5.540	9.380	8.602	6.055	9.362	9.125	3.708	6.496	7.173	5.617	5.769	7.162
	Max	56.621	61.521	58.625	59.097	56.189	52.518	54.561	60.605	51.887	75.813	71.767	70.853	76.578	73.960	72.378
LF/HF	Mean	1.752	1.409	1.377	1.124	1.333	1.323	1.055	1.356	1.311	1.150	1.411	1.304	1.001	1.620	1.505
	SD	1.814	0.930	0.970	1.189	0.866	0.941	0.979	1.008	0.891	1.368	1.273	0.990	1.147	2.278	1.805
	Min	0.275	0.235	0.259	0.229	0.263	0.297	0.267	0.262	0.300	0.178	0.282	0.301	0.214	0.179	0.239
	Max	9.767	4.427	4.910	6.683	4.145	4.720	5.167	4.742	3.645	7.094	5.382	4.439	6.256	12.296	9.779

Table 7 Absolute errors (mean, standard deviation, minimum, and maximum), median of signed errors, Pearson coefficients, and spearman's for five different durations of videos for  $x_1(t)$  and  $y_S(t)$  using M1 and M2.

Parameter	Metrics	5 min		4 min		3 min		2 min		1 min	
		$x_1(t)$ ROI	$y_S(t)$ ICA	$x_1(t)$ ROI	$y_S(t)$ ICA	$x_1(t)$ ROI	$y_S(t)$ ICA	$x_1(t)$ ROI	$y_S(t)$ ICA	$x_1(t)$ ROI	$y_S(t)$ ICA
NN	MAE (ms)	3.812	3.822	3.742	3.656	3.872	4.204	4.689	4.936	5.737	5.543
	SD <sub>AE</sub> (ms)	4.174	4.076	4.254	4.130	4.551	4.520	5.178	4.708	6.012	5.559
	AE <sub>Min</sub> (ms)	0.332	1.210	0.028	0.190	0.063	0.535	0.000	0.221	0.150	0.150
	AE <sub>Max</sub> (ms)	24.311	24.424	24.467	24.496	26.318	26.857	27.621	27.317	29.123	28.371
	Me (ms)	2.193	3.157	2.173	2.405	2.525	2.947	2.459	3.582	2.174	2.130
	$r_p$	0.999	1.000	0.999	0.999	0.999	0.999	0.999	0.999	0.998	0.999
	$\rho_s$	0.998	0.998	0.998	0.998	0.999	1.000	0.997	0.996	0.995	0.994
SDNN	MAE (ms)	5.643	5.853	5.951	5.331	6.563	5.855	6.568	6.466	7.324	6.898
	SD <sub>AE</sub> (ms)	3.197	3.562	3.204	3.057	3.429	3.746	4.478	5.005	4.751	5.191
	AE <sub>Min</sub> (ms)	0.028	1.499	0.848	0.083	0.340	0.329	0.864	0.061	0.026	0.360
	AE <sub>Max</sub> (ms)	10.604	17.355	12.313	12.301	12.866	13.929	16.380	19.619	17.744	21.812
	Me (ms)	4.202	3.867	4.526	4.265	4.553	4.122	3.812	3.634	4.280	3.810
	$r_p$	0.960	0.961	0.962	0.976	0.958	0.963	0.933	0.921	0.913	0.893
	$\rho_s$	0.903	0.928	0.884	0.956	0.895	0.944	0.920	0.911	0.883	0.860
LF	MAE (%)	3.303	2.775	3.577	3.516	4.185	3.915	3.983	4.005	6.204	7.059
	SD <sub>AE</sub> (%)	2.898	2.341	2.970	2.727	3.491	3.056	3.305	2.863	6.183	6.027
	AE <sub>Min</sub> (%)	0.068	0.262	0.290	0.024	0.032	0.005	0.037	0.236	0.097	0.381
	AE <sub>Max</sub> (%)	11.261	8.361	14.605	10.131	13.424	13.275	11.414	13.922	20.921	22.453
	Me (%)	0.783	0.758	2.397	2.529	2.537	2.626	1.893	2.600	2.302	3.540
	$r_p$	0.878	0.916	0.888	0.906	0.901	0.926	0.904	0.917	0.874	0.857
	$\rho_s$	0.813	0.895	0.883	0.890	0.887	0.939	0.864	0.873	0.833	0.730
HF	MAE (%)	6.557	4.372	6.172	5.398	7.290	6.620	8.000	8.351	11.159	11.359
	SD <sub>AE</sub> (%)	5.324	4.684	6.054	4.027	6.312	4.876	7.971	6.399	9.273	8.832
	AE <sub>Min</sub> (%)	0.353	0.592	0.180	0.825	0.398	0.012	0.099	0.243	0.480	0.431
	AE <sub>Max</sub> (%)	20.238	18.450	23.394	16.177	23.250	17.098	33.554	28.548	30.003	34.608
	Me (%)	3.324	1.436	-3.002	-4.042	-3.046	-4.660	-2.459	-6.187	-4.262	-8.664
	$r_p$	0.786	0.885	0.828	0.907	0.782	0.854	0.811	0.845	0.721	0.729
	$\rho_s$	0.779	0.884	0.810	0.889	0.758	0.814	0.690	0.713	0.646	0.705
LF/HF	MAE	0.657	0.482	0.416	0.361	0.481	0.409	0.543	0.428	0.832	0.821
	SD <sub>AE</sub>	0.994	0.927	0.543	0.413	0.634	0.446	0.874	0.547	1.871	1.484
	AE <sub>Min</sub>	0.007	0.002	0.008	0.003	0.018	0.020	0.003	0.016	0.004	0.019
	AE <sub>Max</sub>	5.340	4.857	2.537	1.963	3.160	1.576	3.541	2.656	9.832	7.315
	Me	-0.187	-0.046	0.130	0.208	0.173	0.183	0.112	0.175	0.157	0.249
	$r_p$	0.842	0.931	0.843	0.909	0.723	0.830	0.716	0.881	0.513	0.468
	$\rho_s$	0.779	0.876	0.826	0.915	0.863	0.902	0.833	0.879	0.752	0.765

### 3.5. Discussion

In this chapter we analysed the use of a monochrome video camera to estimate the HR and the HRV from the plethysmographic signal obtained from the human face recordings. Videos from 30 subjects were acquired in rest conditions allowing limited natural movements. An artificial white light was used and the camera input was band-pass filtered so as to match the light absorption peaks of haemoglobin and enhance the BVP contribution to the image grey level. We adopted direct BVP estimation via temporal band-pass filtering (method M1) and BVP peak detection applied to single ROI signals. In addition, ICA pre-processing to extract BVP signal from multiple ROIs (method M2) was implemented. Standard HR and HRV measurements derived from videos were compared with the corresponding descriptors obtained from ECG signals recorded simultaneously to videos.

We analysed both mean absolute errors between video and ECG in three different regions of the face (forehead and each cheek). For completeness, the Pearson and Spearman correlation coefficients were computed. In all cases, the ROI on the forehead provided smallest errors and highest correlations. This seems consistent with a reduced impact of subject motion. Fusion of the three ROIs signals via ICA pre-processing led to substantially reduced errors and improved correlation in all frequency domain descriptors, whereas time domain descriptors exhibit a behaviour similar to the ones from the ROI on the forehead.

At the original video rate of 133 fps method M1 produced acceptable results for the forehead ROI (MAEs were 3.812 ms for NN, 5.643 ms for SDNN, 3.303 % for LF, 6.556 % for HF, and 0.657 for LF/HF).

Anyway, M2 provided data more consistent with ECG estimation as made evident by frequency domain parameters (MAEs were 3.822 ms for NN, 5.853 ms for SDNN, 2.775 % for LF, 4.372 % for HF, and 0.482 for LF/HF).

When the video signal is down-sampled from the native 133 fps to 66 fps and 33 fps estimates of the HRV descriptors worsen, whereas the HR does not look significantly affected. This finding is in line with the expected need of an accurate detection of the BVP pulses for the HRV analysis, a requirement that can be significantly relaxed for the HR assessment. In particular, the LF/HF parameter seems to be the most affected. It gets worse with the decrease of the sampling rate. In the frequency domain, the contribution of the ICA pre-processing is significant at low sampling rates.

As to the impact of video duration, the HR, as expected, does not exhibit a particular sensitivity and both M1 and M2 produce constant and similar results. In the case of the HRV, method M1 looks less robust than method M2 and discrepancies arise between iPPG and ECG for very short recordings (1 minute). In general, the MAEs increase when the sampling rate and the acquisition time decrease.

To summarize, both M1 and M2 can be expected to allow a very confident estimation of the HR. As concerns the HRV descriptors, the ICA pre-processing enables a more reliable estimation in all considered experimental conditions.

## **Chapter 4. Unobtrusive monitoring of stress indicators at rest**

In this chapter we focus on the assessment of resting status of two groups of voluntaries with different age ranges. Evaluating the status in the absence of external stimuli is crucial to assess and understand individual stimuli responses.

To this end we focused on the results obtained by the integration of simple unobtrusive devices monitoring skin blood volume pulses, respiratory rate, and galvanic skin response. Data were recorded in resting conditions and analysed as described in the following sections.

### **4.1. Materials and methods**

#### **4.1.1. Study's participants and experimental setup**

The data of twelve healthy volunteers were analysed for this chapter. The sample was split into two groups: group I was composed by six voluntaries (4 females and 2 males) with age between 25 and 35 (mean 29, min. 25, max. 35); group II was composed by six voluntaries (4 females and 2 males) with age between 55 and 65 (mean 60.16, min. 53, max. 63).

The subject's signals measured were HR, HRV, RR, and GSR. The devices included in the experimental setup of this study were the camera, the BioHarness 3 Zephyr, and the Shimmer3 GSR. The experimental setup was deeply described in section 2.2. In this study we analysed the data acquired during the rest state only.

#### **4.1.2. Data analysis**

The video signal analysis was performed through M1 procedure as described in section 3.3.5. Concerning the respiratory rate, the values of the median and the interquartile range were calculated. For GSR, the interfering main frequency (50 Hz) was removed from the signal by a notch filter. The average slope of the descending portion of the GSR signal was estimated by linear regression (see Figure 11).

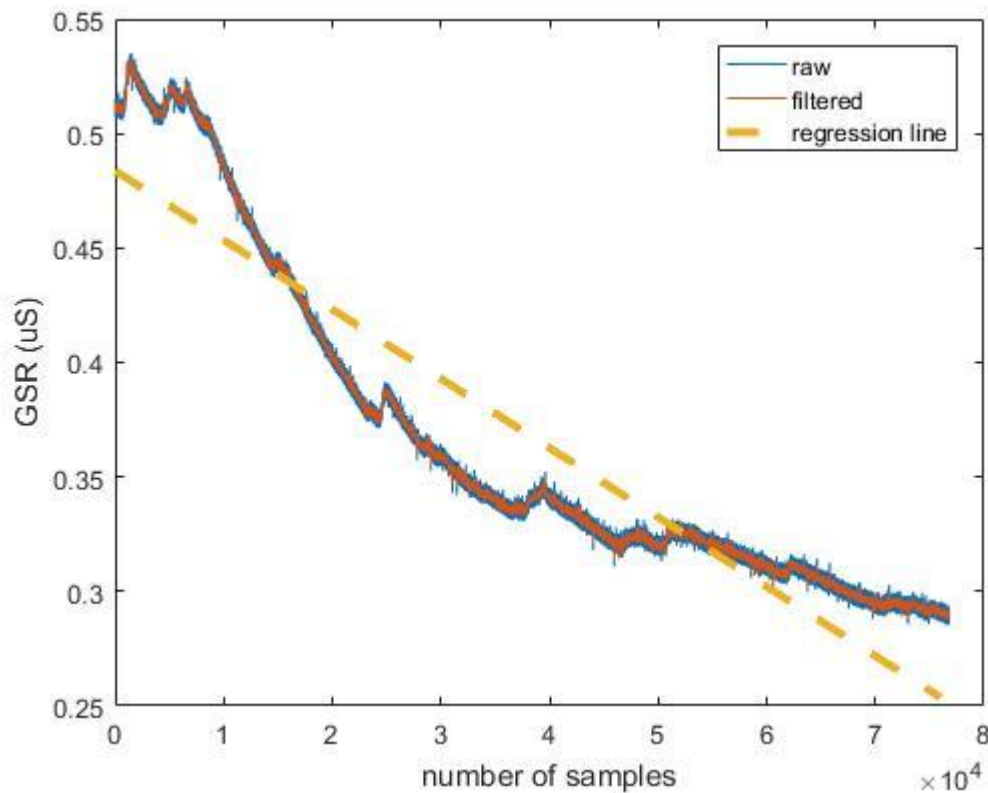


Figure 11 A typical GSR record. The raw and the filtered signals of the skin conductivity of a subject during the resting state are shown. The regression line (obtained by the linear regression) is the average slope of the signal.

## 4.2. Results

Collected data were analysed for the two groups separately. In Table 8, the median and the interquartile range of all the data are reported.

The median value (and interquartile range) of NN as calculated from the video signal is 823.24 ms (171.19 ms), and 849.98 ms (50.87 ms) for group I and group II respectively. This parameter is smaller in group I than in group II. The median value (and interquartile range) of RR as measured through BioHarness 3 Zephyr is 14.92 bpm (3.86 bpm), and 13.29 bpm (0.66 bpm) for group I and group II respectively. This parameter is smaller in group II than in group I. The variability of NN and RR in group II is lower than in group I. As shown in Figure 11, the skin conductivity decreases when the subject relaxes. The median (interquartile range) of the slope of the skin conductivity trend is  $-4.15 \times 10^{-4}$  ( $1.16 \times 10^{-3}$ ), and  $-1.62 \times 10^{-3}$  ( $3.08 \times 10^{-3}$ ) for group I and group II respectively. This suggests that skin conductivity of group II decreases faster than that one of group I.

Table 8 Results summary. For each group, the median and the interquartile range, in brackets, of HR and HRV parameters (NN, SDNN, LF, HF, LF/HF), respiratory rate, and the slope of the skin conductivity trend (GSR<sub>SLOPE</sub>).

Parameter	Group I	Group II
NN <sub>VIDEO</sub> (ms)	823.24 (171.19)	849.98 (50.87)
SDNN <sub>VIDEO</sub> (ms)	86.98 (52.02)	90.01 (48.32)
LF <sub>VIDEO</sub> (normalized unit)	56.54 (47.44)	54.06 (9.92)
HF <sub>VIDEO</sub> (normalized unit)	43.46 (44.44)	45.94 (9.92)
LF <sub>VIDEO</sub> /HF <sub>VIDEO</sub>	1.32 (1.73)	1.18 (0.52)
RR (bpm)	14.92 (3.86)	13.29 (0.66)
GSR <sub>SLOPE</sub>	$-4.15 \times 10^{-4}$ ( $1.16 \times 10^{-3}$ )	$-1.62 \times 10^{-3}$ ( $3.08 \times 10^{-3}$ )

### 4.3. Discussion

In this chapter we reported the results obtained by the integration of simple unobtrusive devices monitoring skin blood volume pulses, respiratory rate, and galvanic skin response. Data from two groups of volunteers with different age ranges were recorded in resting state. Evaluating the status in the absence of external stimuli is crucial to assess and understand individual stimuli responses.

The results indicate that the variability of NN and RR is smaller in group II than in group I. The skin conductance decreases when the subject relaxes and decreases faster in group II than in group I.

The study reported in this chapter allowed us to validate the experimental setup. The latter turned out robust and reliable for unobtrusively monitoring the signals taken into account, permitting us to extract relevant parameters from two groups of volunteers.

## **Chapter 5. Unobtrusive monitoring of EEG signals during mild cognitive activation**

In this chapter we report on the acquisition and analysis of EEG signals during mild cognitive activation. EEG were acquired simultaneously to skin blood volume pulses, and the respiratory rate by the set of unobtrusive devices described in section 2.2. Two different condition were compared: rest state and mild cognitive activation.

### **5.1. Materials and methods**

#### **5.1.1. Study's participants and experimental setup**

The data of seven healthy participants were analyzed for this study. The sample was composed of 5 females and 2 males with mean age of 41.72 years, range from 25 to 62.

The subject's signals measured were HR, HRV, RR, and EEG. The signals were recorded in rest state for 5 minutes and during a mild cognitive activation induced by two different versions of the Stroop color word test, a congruent one and an incongruent one (Test A and Test B). The devices included in the experimental setup of this study were the ECG system, the camera, the BioHarness 3 Zephyr, and the Muse 2 Headband. These devices and the experimental setup were deeply described in section 2.2 and 2.3.

#### **5.1.2. Data analysis**

The ECG signal analysis was performed as described in section 3.3.3. The video signal was analyzed through M1 (described in section 3.3.5). Concerning the respiratory rate, the values of the median and the interquartile range were calculated.

The Muse 2 Headband is a simple electroencephalography device. It has three reference sensors on the forehead, two channels on the left (antero-frontal AF7 and temporo-parietal TP9) and two on the right (antero-frontal AF8 and temporo-parietal TP10). AF7 and AF8 are forehead sensors and TP9 and TP10 are Smart Sense conductive rubber ear sensors. The band waves were calculated, these are the delta, theta, alpha, and beta. The delta wave is the frequency range from 0.5 Hz to 4 Hz; the theta wave is the frequency range from 4 Hz to 7 Hz; the alpha wave is the frequency from 7 Hz to 13 Hz, the beta wave is the frequency from 14 Hz to 30 Hz. The features extracted from the EEG [54] through a custom software written in C include the normalized band power for each channel, that is calculated as:



$$100 \frac{\int_{f_1}^{f_2} P}{P_t} \quad (4)$$

where  $f_1$  and  $f_2$  are the lower and the upper frequency of each band respectively (for example for the alpha band  $f_1$  is 7 Hz and  $f_2$  is 13 Hz),  $p$  is the band power and  $P_t$  is the total power. The second feature extracted was the each band's (delta, theta, alpha, beta) power asymmetry (delta power asymmetry-DPA, theta power asymmetry-TPA, alpha power asymmetry-APA, beta power asymmetry-BPA), that is calculated as:

$$\frac{EEG_{bandpowerR} - EEG_{bandpowerL}}{EEG_{bandpowerR} + EEG_{bandpowerL}} \quad (5)$$

where  $EEG_{bandpowerR}$  and  $EEG_{bandpowerL}$  are respectively the band power measured from the sensor on the right side of the head and on the left side of the head.

The four signals were jointly processed through Independent Component Analysis as implemented in FastICA algorithm [119]; four new sequences were so obtained, each representing a different signal contribution. One of these signals was identified as the one with eyes blinking and, thanks to the help of an expert, it was removed. The remaining signals were analysed through the inverse FastICA in order to obtain four signals again (Figure 12). The analysis was performed by power spectrum density estimated by the Lomb-Scargle periodogram [114].

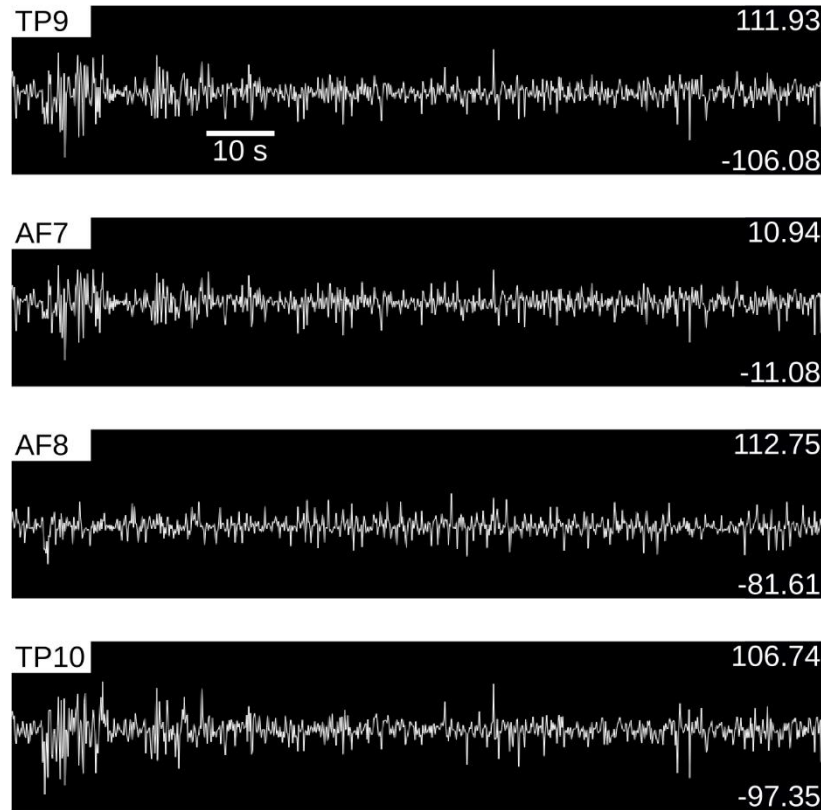


Figure 12 The EEG signals after ICA processing.

## 5.2. Results

In Table 9 we reported the median and the interquartile range of each parameter as observed in the study sample. Concerning the ECG results, the median value (and interquartile range) of NN is 926.98 ms (128.78 ms), 820.66 ms (191.41 ms), 848.98 ms (196.64 ms) in the three conditions respectively (rest, Test A, and Test B). Concerning the video results, the median value (and interquartile range) of NN is 851.74 ms (144.62 ms), 822.67 ms (272.95 ms), and 796.10 ms (143.68 ms), in the three conditions respectively (rest, Test A, and Test B). The median value of RR is 13.08 bpm, 15.17 bpm, and 16.51 bpm, in the three conditions respectively (rest, Test A, and Test B). The most relevant feature extracted from the EEG is the power asymmetry from the left and right normalized alpha band power of AF7 and AF8 sensors between rest and Test A, and of TP9 and TP10 between rest and Test B. The APA of AF7 and AF8 was significantly smaller during Test A (median value = -0.35) than in rest (median value = 0.27) with p-value = 0.01. The APA of TP9 and TP10 was significantly smaller during Test B (median value = 0.04) than in rest (median value = 0.10) with p-value = 0.01.

Table 9 For each condition, the median and the interquartile range, in brackets, of the all parameters extracted.

Parameter	Rest	Test A	Test B
NN <sub>ECG</sub> (ms)	926.98 (128.78)	820.66 (191.41)	848.98 (196.64)
SDNN <sub>ECG</sub> (ms)	47.94 (42.80)	42.40 (29.77)	55.17 (36)
LF <sub>ECG</sub> (normalized unit)	79.09 (24.80)	54.86 (19.16)	52.07 (34.23)
HF <sub>ECG</sub> (normalized unit)	20.91 (24.80)	45.14 (19.16)	47.93 (34.23)
LF <sub>ECG</sub> /HF <sub>ECG</sub>	3.78 (2.93)	1.22 (1.18)	1.09 (1.63)
NN <sub>VIDEO</sub> (ms)	851.74 (144.62)	822.67 (272.95)	796.10 (143.68)
RR (bpm)	13.08 (2.66)	15.17 (4.22)	16.51 (3.18)
APA (AF)	0.27 (0.65)	-0.35 (0.28)	/
APA (TP)	0.10 (0.33)	/	0.04 (0.14)

## 5.3. Discussion

We have reported the results obtained by the integration of simple unobtrusive devices monitoring skin blood volume pulses, respiratory rate, and EEG signal. The subjects involved in the study underwent a mild cognitive activation.

During the stress condition, the heart rate and the respiratory rate increased, and changes of the brain activity was observed. In particular the last factor indicates that the

right alpha power was reduced to a greater extent than the left alpha power in a stress situation, which is consistent with the physiological assumptions (i.e., enhanced activation occurred in the right hemisphere, which shows negative emotions).

The results obtained demonstrated that the devices and the procedure involved in this study allow detecting differences in a mild stress condition using unobtrusive measurements of psychophysical parameters.

## **Chapter 6. Unobtrusive monitoring of stress indicators during mild cognitive activation**

In this chapter we report the analysis of the “Mild Cognitive Activation” (MCA) dataset collected from the volunteers enrolled during my work.

We focused on the results obtained by the integration of simple unobtrusive devices monitoring skin blood volume pulses, respiratory rate, and galvanic skin response. Data were recorded both in rest state and during a mild cognitive activation.

### **6.1. Materials and methods**

#### **6.1.1. Study’s participants and experimental setup**

The data from ten healthy volunteers were analysed for this study. The sample was composed of 7 females and 3 males with mean age of 45 years, range from 25 to 62. Two of these subjects (namely subject 08 and 09) were considered non-naïve subjects, as they had previous experience with the Stroop colour word test.

The subject’s signals measured were HR, HRV, RR, and GSR. The signals were recorded in rest state for 5 minutes and during a mild cognitive activation induced by two different versions of the Stroop colour word test, a congruent one and an incongruent one (Test A and Test B). The devices included in the experimental setup of this study were the ECG system, the camera, the BioHarness 3 Zephyr, and the Shimmer3 GSR. These devices and the experimental setup were deeply described in section 2.2 and 2.3.

#### **6.1.2. Data analysis**

The ECG signal analysis was performed as described in section 3.3.3. The video signal was analysed through the algorithm M1 (described in section 3.3.5). Concerning the NN values both of ECG and video signal, the median and the interquartile range were calculated. Regarding the RR, the values of the median, the interquartile range, the minimum, and the maximum were calculated. As to the galvanic skin response, the interfering mains frequency (50 Hz) was removed from the signal by a notch filter. In order to calculate the GSR parameters, the phasic component was extracted. To this end and whereas the GSR signal is often sampled at a much higher sampling rate than actually required, the GSR signal was down sampled to 10 Hz. Then it was filtered through a mean filter spanning a 4 s window and a median filter spanning 8 s. Then the signal from the median filter was subtracted to the one from the mean filter (Figure 13). The number of peaks, the maximum peak amplitude, and the median peak value were calculated.

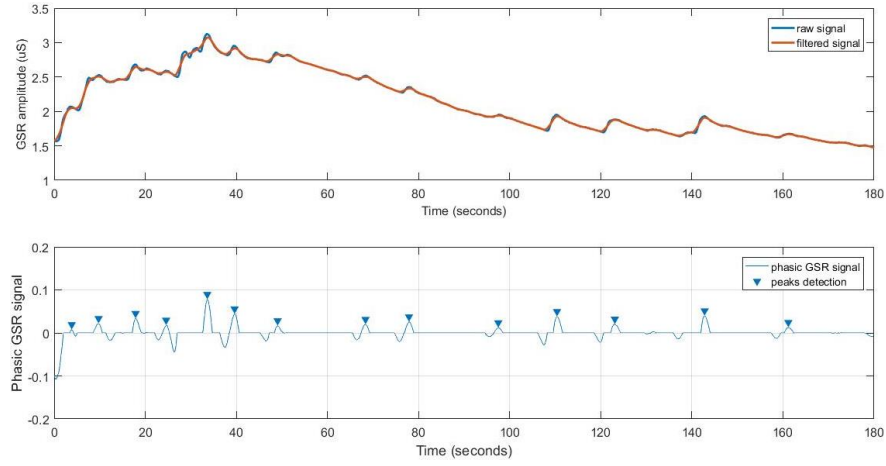


Figure 13 An example of GSR signal. Upward the raw and the filtered signal, below the phasic signal and the peaks detection.

## 6.2. Results

Figure 14 shows the box plot of the NN value calculated by the ECG signal (left) and for the video camera (right), for the three conditions (rest, Test A, and Test B). As to the ECG results, the median value of the NN (and the interquartile range) is 907.51 ms (102.08 ms), 825.61 ms (165.58 ms), 815.8 ms (161.82 ms), respectively in rest, Test A, and Test B. Concerning the video camera results, the median value of the NN (and the interquartile range) is 858.17 ms (121.21 ms), 784.29 ms (129.5 ms), 778.60 ms (170.13 ms), respectively in rest, Test A, and Test B. The NN value both for ECG signal and video camera signal decreases in the stress condition. The Pearson correlation coefficient between the NN value of the ECG and the video signal is equal to 0.92, 0.88 and 0.85, respectively for rest, Test A, and Test B ( $p$ -value < 0.007).

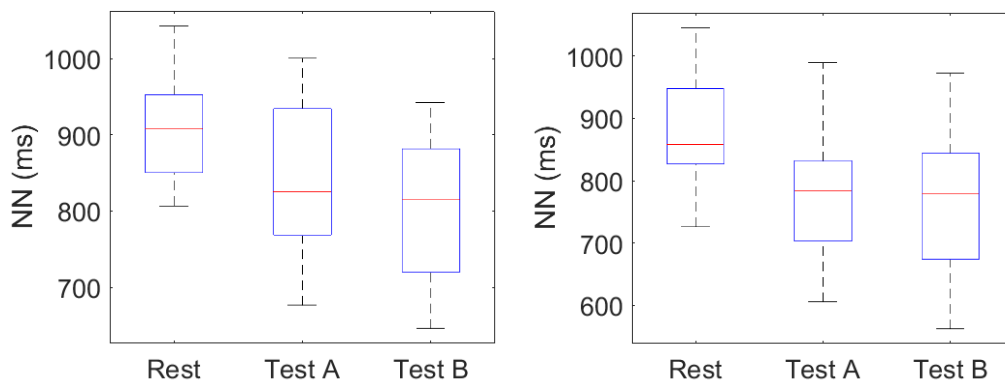


Figure 14 Box plot of the NN value of the ECG signal (left) and the video signal (right) in the three experimental conditions.

In Figure 15 is showed the box plot of the of the RR value for the three experimental conditions (rest, Test A, and Test B). The median value (and the interquartile range) of the RR is 13.5 bpm (2.69 bpm), 16.43 bpm (3.24 bpm), 16.67 bpm (1.91 bpm), respectively in rest, Test A, and Test B. The RR value increases in the stress condition.

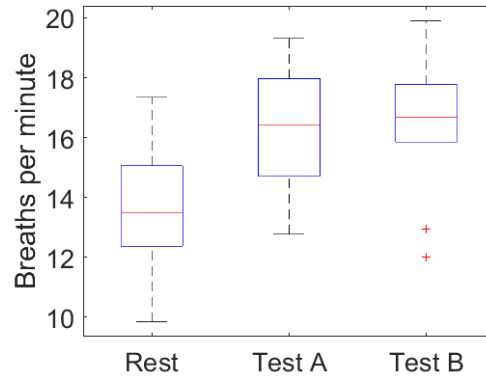


Figure 15 The box plot of the RR value in the three experimental conditions.

In Table 10 we reported, for each subject, the median peak value and the number of peaks of the GSR signal calculated in the three experimental conditions (rest, Test A, and Test B). In general, these values increase in the stress condition. For subject 08 and 09 we obtained zero peaks from the GSR, in all experimental conditions. It is worth noting that these subjects were not naïve to the Stroop colour word test.

Table 10 Median peak value ( $\mu$ S) and number of peaks of the GSR for the three experimental conditions.

Subject ID	Median peak value ( $\mu$ S)			Number of GSR peaks		
	Rest	Test A	Test B	Rest	Test A	Test B
01	1.49	2.87	2.57	0	9	16
02	0.43	1.90	1.86	0	12	7
03	0.34	1.83	1.96	0	4	14
04	0.48	1.15	1.45	0	9	22
05	0.23	0.99	1.18	0	11	21
06	1.74	5.04	5.22	11	20	25
07	2.73	12.54	15.32	7	11	25
08	0.38	0.39	0.43	0	0	0
09	0.94	1.57	1.98	0	0	0
10	1.98	4.20	4.85	5	15	29

### **6.3. Discussion**

The results in this chapter were obtained by the integration of simple unobtrusive devices monitoring skin blood volume pulses, respiratory rate, and galvanic skin response. The signals were monitoring both in rest state and in a mild stress condition.

During the stress condition, the heart rate, the respiratory rate, and the median peak value increased for eight subjects. No peaks of the galvanic skin response were registered for the two non-naïve subjects for all the three conditions (rest, Test A, and Test B). These two subjects had previous experience with the Stroop colour word test and remained in a rest state also during the mild cognitive activation.

The results are in accordance with literature and, as we have already demonstrated, confirmed that the devices and the method involved in this study allow detecting differences in a mild stress condition through unobtrusive measurements of psychophysical parameters.

## Chapter 7. Data integration

The final objective of this thesis was to develop a model of the wellness state of a person endowed with predictive power. This is the subject of the activities reported in this chapter.

As the development of a model usually requires as much data as possible, we wanted to enrich/integrate the MCA dataset collected from the volunteers enrolled during my work, also to mitigate the impact that the COVID-19 lockdown had on the planned acquisition campaign. To do so, we searched for similar publicly datasets. During this search, particular attention was paid to the data hosted in the Physionet repository [120]. Among these, the “Stress Recognition in Automobile Drivers” (SRAD) dataset from the MIT Media Lab was deemed relevant for the current project and was analysed in detail.

A set of features were extracted from the available signals (including ECG, HR, RR, GSR, EMG acquired both at rest and during driving). The data harmonization and the feature selection were performed using two complementary approaches: sequential forward feature selection (SFFS) and auto-encoder (AE) neural network. Then a cluster analysis was carried out using two complementary methods: k-means clustering and Kohonen self-organizing map (SOM). Subsequently the clustering was validated using the MCA dataset.

### 7.1. Datasets

#### 7.1.1. The MCA dataset description and its analysis

For each subject, the MCA dataset included three blocks of data. The first block was 5 minutes long and contained the data acquired during the rest state. The other two blocks were 2 minutes and 3 minutes long respectively, they contained data acquired during the mild cognitive activation induced by two different versions of the Stroop colour word test (Test A and Test respectively, described in section 2.3) . For our aims, these two blocks were merged into a single 5-minutes window. In this way each record contains two 5-minute blocks. The data analysis is described in section 6.1.2.

#### 7.1.2. The SRAD dataset

Among the analysed repositories, the SRAD dataset from MIT Media Lab, hosted by Physionet.org [120] resulted particularly promising for our purposes. This database contains a collection of multiparametric recordings from 17 healthy volunteers. The data were acquired during driving. Records include two 15 minutes rest periods occurred one at the beginning and one at the end of a driving period. During these periods, the driver sat in the garage with eyes closed and with the car in idle. The rest periods were used to gather



baseline measurements and to create a low stress situation. After the first rest period, drivers drove, for around 50 minutes, on a prescribed route including city streets and highways in and around Boston, Massachusetts.

Four types of physiological sensors were used during the experiment: electrocardiogram, electromyogram, galvanic skin response, and respiration (through chest cavity expansion). The ECG electrodes were placed in a modified lead II configuration. The EMG was placed on the trapezius (shoulder) and has been used as an indicator of emotional stress [121]. The galvanic skin response was measured in two locations: on the palm of the left-hand using electrodes placed on the first and middle finger and on the sole of the left foot using electrodes placed at each end of the arch of the foot. Respiration was measured through chest cavity expansion using an elastic Hall-effect sensor strapped around the driver's diaphragm. Figure 16 shows the general placement of sensors with respect to the automotive system. The ECG was sampled at 496 Hz, the galvanic skin response and respiration sensor were sampled at 31 Hz, and the EMG was sampled at 15.5 Hz after first passing through a 0.5 s averaging filter.



Figure 16 Placement of recording instruments as deployed while collecting the “Stress Recognition in Automobile Drivers” dataset. Picture reproduced from [120].

### 7.1.3. The SRAD dataset analysis

Seven subjects were excluded from the analysis due to the lack of some data, a sub-group of 10 subjects was therefore analysed. Among the signals included in the dataset we considered those in common with our dataset. These were: the ECG, the respiratory wave, and the GSR acquired on the palm of the hand.

The ECG signal was firstly pre-processed so as to derive the related tachogram. QRS complexes were detected by Pan-Tomkins algorithm [112]. The tachogram was then analysed in 5-minute sliding window (with a 5-minute stride) using the same method described in section 3.3.3.

The respiratory rate was extracted from the respiratory wave of the chest belt. Particularly, the respiratory rate (in breaths per minute) was calculated by detecting and counting the signals peaks in sliding windows of 1 minute length (step of 1 s). The GSR was analysed as described in section 6.1.2.

The signals were divided in (non-overlapping) five minutes long windows. As each record in the SRAD dataset contained 15 minutes of rest periods both at the beginning and at the end of acquisition, for each subject, we had six windows representing the resting condition. Data from all other windows represented the driving activation period. For each window, the following parameters were calculated. Concerning the ECG signal, in time domain, we calculated the average time between adjacent normal heartbeats and its standard deviation. Concerning the frequency domain, analysis of the HRV was performed by the power spectrum density estimated by the Lomb-Scargle periodogram and the low frequency, the high frequency, and their ratio were calculated. Concerning the respiratory rate, the median, the interquartile range, the minimum, and the maximum value were calculated. Concerning the galvanic skin response, the number of peaks, the maximum peak amplitude, and the median peak value were calculated.

Table 11 Extracted features and related source signals.

Feature	Source	
	MCA dataset	SRAD dataset
Median RR	Chest belt	Chest belt
RR interquartile range	Chest belt	Chest belt
Minimum RR	Chest belt	Chest belt
Maximum RR	Chest belt	Chest belt
Number of GSR peaks	GSR	GSR
Maximum GSR peak amplitude	GSR	GSR
Median GSR peak amplitude	GSR	GSR
NN	Video iPPG	ECG
SDNN	Video iPPG	ECG
LF	Video iPPG	ECG
HF	Video iPPG	ECG
LF/HF	Video iPPG	ECG

## 7.2. Features selection

The analysis described above led to a large number of parameters to deal with (Table 11). Thus, although the dataset available was expanded through the one found in literature, a feature selection process was deemed necessary to avoid the curse of dimensionality before proceeding to the actual definition of a model.

To this end we exploited a sequential forward feature selection method which can also provide indications about the relevance of different features at discriminating rest from non-rest conditions.

Simple feature selection is based on the use of a possibly optimal subset of available features, and may hinder the correlation among features. During the features selection process, dimensionality reduction is usually achieved by completely discarding some dimensions, which inevitably leads to loss of information. However, sample data in a high dimensional space generally cannot diffuse uniformly in the whole space; they actually lie in a low dimensional manifold embedded in the high dimensional space, the dimensionality of the manifold being called the intrinsic dimensionality of the data [122].

To investigate this aspect, we explored an alternative approach based on AEs. They provide a sort of non-linear generalization of principal component analysis [123]. AEs are unsupervised neural networks that learn to output a reconstruction of the input. The simplest multilayer architecture for building an auto-encoder is to constrain the number of nodes of the hidden layer(s), limiting the amount of information flowing through the network. AEs are commonly used for data dimensionality reduction being able to capture the intrinsic data dimensionality minimizing the information loss and providing an optimal set of synthetic features.

### 7.2.1. Sequential forward feature selection

As a preliminary analysis, we explored a SFFS method. In this method, the best single feature is selected using some criterion function. Then, pairs of features are formed using one of the remaining features and this best feature, and the best pair is selected. Next, triplets of features are formed using one of the remaining features and these two best features, and the best triplet is selected. This procedure continues until a predefined stop criterion is met. This method was selected as a simple feature selection method and was considered the perfect balance between efficacy (generally higher than filter methods) and computational cost (generally lower than exhaustive) [124].

The process was implemented by the MATLAB `sequentialfs` function which selects a subset of features from the data matrix the best predict the data by sequentially selecting features until there is no improvement in prediction. Prediction was implemented by

standard linear discriminant analysis (LDA) [125]. LDA is a well-known supervised method used to find a linear combination of features that characterises or separates objects from two or more classes. For each candidate feature subset, sequentialfs performs 10-fold cross-validation by repeatedly training and testing a model with different training and test subsets.

As sequentialfs randomly splits the initial dataset to implement 10-fold cross-validation, the feature selection process can yield to different results depending on the run. This is true for both the number of features selected and which features are selected. In order to mitigate this effect, sequentialfs was run 1000 times. For each run, we recorded the features selected by the procedure and the order in which they were selected. In addition, we set up a scoring system to properly weigh the relevance of the features. Specifically, at each run, every selected feature got a score equal to its position in the selection process (e.g. 1 for the first one, 3 for the third one). The features that were not selected were given a score of 12. The process was repeated for every run and the scores for each run were summed up. By doing so, a final score of 1000 would indicate a feature selected as the first one in all runs. A feature with a final score of 12000 would be a feature never selected by the method.

The method was trained using the SRAD dataset, whilst MCA dataset was used as independent test set.

### **7.2.2. Auto-Encoder neural network**

The aim of an AE is to learn a representation (encoding) for a set of data, typically for dimensionality reduction, by training the network to ignore “noise”. Along with the reduction side, a reconstructing side is learnt, where the AE tries to generate from the reduced encoding a representation as close as possible to its original input. We used the dimensionality reduction property of the AE that allow to maps data from a large feature space to a small feature space.

An auto-encoder neural network was applied on the SRDA dataset. This was designed as a neural network with a single hidden layer. To build the AE, a 5-fold cross-validation was performed by repeatedly training and testing a model with different train and test subsets of the data. In order to find the best trade-off between accuracy and data dimensionality, the size of the hidden layer was changed from 1 to 12. Using each set of the new features generated by the AE, an LDA classifier was trained in order to predict the driver stress level. This process generated 12 models of data classification, each one based on a number of features varying between 1 and 12.

Finally, the 12 models were applied on the MCA dataset in order to validate the results against new data.

### 7.3. Detection of activation status

Recognition of activation status was firstly investigated by k-means clustering. K-means clustering is a method of vector quantization, originally from signal processing, that aims to partition  $n$  observations into  $k$  clusters in which each observation belongs to the cluster with the nearest mean (cluster centres or cluster centroid), serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. K-means clustering minimizes within-cluster variances (squared Euclidean distances).

The k-means clustering is a widely useful data analysis tool, however it requires the user to set the number of clusters a-priori. We resorted to use the Kohonen self-organizing map, which exploits self-organization mechanisms and has the ability to build accurate, but low dimensional, topology preserving-maps [125]. This means that neighbouring inputs vectors tend to excite neighbouring units in the map. Topology preservation is ensured by a competitive learning, network parameters being adapted according to a neighbourhood function.

#### 7.3.1. Activation representation by k-means clustering

In this work we used the MATLAB implementation of the k-means algorithm. The data were grouped into two clusters: one being expected to represent the rest condition and the other one the driving period. The latter was associated with a higher stress level.

Afterwards, the data from the MCA dataset were used to validate these results against new data, not used for

clustering. As described above, the MCA dataset was composed, for each subject, of two windows of data. Each data point, corresponding to one window, was associated to one cluster or the other, depending the one having the closest centroid.

As an additional investigation, we tested the sensitivity of the clustering process to the cluster's initialization (i.e. their initial value). To do so, we repeated the clustering process 1000 times, forcing the clusters initialization across the whole feature space.

#### 7.3.2. Activation representation by SOM maps

The Kohonen self-organizing map method uses unsupervised learning to produce a low-dimensional (typically two-dimensional), discretised representation of the input space of the training samples, called a map, and is therefore a method to do dimensionality reduction. Self-organizing maps differ from other artificial neural networks as they apply competitive learning as opposed to error-correction learning (such as back propagation with gradient descent), and in the sense that they use a neighbourhood function to preserve the topological properties of the input space.

Like most artificial neural networks, SOMs operate in two modes: training and mapping. “Training” builds the map using input examples (a competitive process, also called vector quantization), while “mapping” automatically classifies a new input vector.

The visible part of a self-organizing map is the map space, which consists of components called nodes or neurons. The map space is defined beforehand, usually as a finite two-dimensional region where nodes are arranged in a regular hexagonal or rectangular grid. Each node is associated with a “weight” vector, which is a position in the input space; that is, it has the same dimension as each input vector. While nodes in the map space stay fixed, training consists in moving weight vectors toward the input data (reducing a distance metric) without spoiling the topology induced from the map space. Thus, the self-organizing map describes a mapping from a higher-dimensional input space to a lower-dimensional map space. Once trained, the map can classify a vector from the input space by finding the node with the closest (smallest distance metric) weight vector to the input space vector.

The SRAD dataset was used to train the clustering algorithms. Using the Matlab Neural Network package, we have analysed 2D maps of varying dimensions, from 3 x 3 to 6 x 6 units. Networks were trained with the SOM batch algorithm [126] using the SRAD dataset.

The obtained maps were tested using the MCA data.

## 7.4. Results

### 7.4.1. Sequential forward feature selection

Results of the feature selection process are summarized in plot the accuracy of LDA classifier obtained on the SRAD dataset with different numbers of features according to SFFS ranking. A similar plot is also given for MCA as test set. For SRAD the accuracy starts at about 93 % for a single feature rising to about 98 % with two features with no further significant changes using the remaining features. For the MCA data we observed a rapid increase to about 90 % with two features, fluctuations being present when using more than six features.

Table 12 where features are ranked according to the sequential forward feature selection score, selection frequency is also reported. In particular, the median RR was selected in all runs, and is constantly the most relevant single feature. This suggests that it carries a significant piece of information, irrespectively of how the data is split between training and test set. The number of GSR peaks totalized the second score in the process, being selected in almost 99 % of the repetitions.

After the first two features, we observed a net drop in the score. Indeed, LF/HF (the feature with the third best score, i.e. > 9200) was selected in 34 % of the runs only. Similar results were observed for the RR interquartile range and NN. Finally, all other features had scores that are very close to the maximum (> 10000) to end up with the last one (the

maximum RR) never being selected by the process. Table 13 and Table 14 show values for the two most relevant features for all subjects of SRAD and MCA, respectively, used in this work.

In Figure 17 Accuracy of SFFS process. On the left: accuracy of LDA classifier evaluated on the SRAD dataset varying the number of feature according to SFFS ranking. On the right: accuracy evaluated on MCA. Figure 17 we plot the accuracy of LDA classifier obtained on the SRAD dataset with different numbers of features according to SFFS ranking. A similar plot is also given for MCA as test set. For SRAD the accuracy starts at about 93 % for a single feature rising to about 98 % with two features with no further significant changes using the remaining features. For the MCA data we observed a rapid increase to about 90 % with two features, fluctuations being present when using more than six features.

Table 12 The score and the selection frequency (SF) obtained by each features.

Features	Score	SF (%)
Median RR	1000	100
Number of GSR peaks	2146	98.7
LF/HF	9220	34.2
RR Interquartile range	9571	30.9
NN	9842	26.2
SDNN	10958	11.8
LF	11501	7.0
Minimum RR	11700	4.0
Maximum GSR peak amplitude	11850	1.7
HF	11974	0.5
Median GSR peak amplitude	11991	0.1
Maximum RR	12000	0

Table 13 SRAD database: median RR and number of GSR peaks both at rest and during driving. Interquartile range is reported in brackets.

Subject ID	RR (bpm)		Number of GSR peaks	
	Rest	Driving	Rest	Driving
01	13.5 (3)	24.5 (3)	8 (6)	22.5 (5)

02	15.5 (3)	23 (4.5)	0 (0)	21 (8.5)
03	14 (2)	20 (2)	10.5 (2)	25 (5)
04	12 (2)	20.5 (2)	0 (2.25)	17 (7)
05	15 (0.75)	21 (2)	0 (0)	15.5 (3)
06	13 (0)	20 (3.75)	3 (0)	18 (6.75)
07	16 (1)	21 (3.25)	0 (0)	21 (5.5)
08	18 (0)	24 (1.625)	6 (6)	26 (10.25)
09	14 (3)	17 (2)	5 (12)	18 (1.5)
10	18 (0.75)	24.5 (2)	15 (8.25)	31 (4)

Table 14 MCA database: median RR with interquartile range (in brackets) together with the number of GSR peaks both at rest and during SCWT. No interquartile range is available for number of GSR peaks as we only have a single block.

Subject ID	RR (bpm)		Number of GSR peaks	
	Rest	SCWT	Rest	SCWT
01	14 (1)	17.3 (2.7)	0	25
02	16 (3)	17 (1.5)	0	19
03	16 (2)	19.5 (2)	0	18
04	13 (2)	13.5 (1)	0	32
05	11 (4)	18 (3.5)	0	25
06	10 (2)	14.5 (4.5)	11	45
07	17 (1)	19.5 (1)	7	36
08	15 (1)	17 (0)	0	0
09	13 (1)	15.5 (1)	0	0
10	13 (1)	13 (2)	5	44



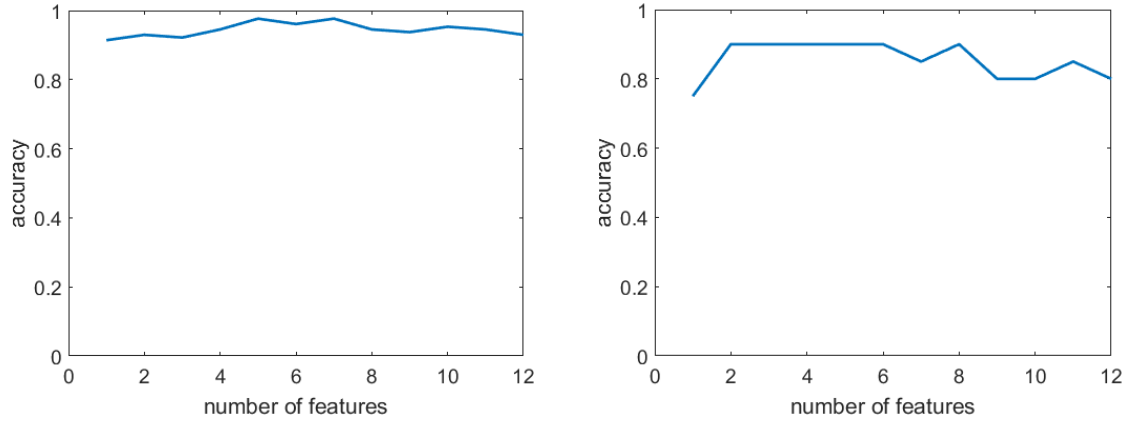


Figure 17 Accuracy of SFFS process. On the left: accuracy of LDA classifier evaluated on the SRAD dataset varying the number of feature according to SFFS ranking. On the right: accuracy evaluated on MCA.

#### 7.4.2. Auto-Encoder neural network

In general, the observed accuracy of AE features resulted high ( $> 90\%$ ), with the worst performance (about  $93\%$ ) obtained when a single feature was available (Figure 18). When two (or more) features were employed, the accuracy fluctuated around  $93 - 95\%$ . Therefore using more than two AE features did not significantly improve discrimination capabilities. Indeed, a maximal accuracy of  $95\%$  was already met using two features.

When each of the 12 models were tested on the MCA dataset, results were found to be affected by a larger variability (Figure 18). With a single feature, the accuracy is  $50\%$  (chance level). However, also in this case, using two features produces a substantial accuracy boosting. This reached its maximum with three features (at around  $90\%$ ). The inclusion of additional features results in accuracy fluctuation about a lower value ( $55\%$ ). To sum up, both SFFS and AE method supports the finding that two features may be sufficient to reliably recognize activation statuses.

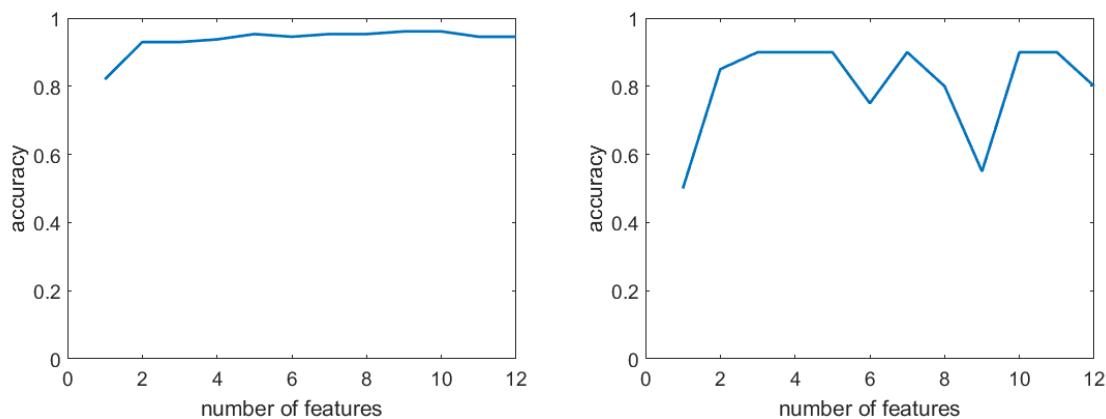


Figure 18 Accuracy of AE process. On the left: accuracy of LDA classifier evaluated on the SRAD dataset with different auto-encoding dimensionality. On the right: accuracy evaluated on MCA.

### 7.4.3. K-means clustering

According to the results of the selection process, the best features were identified as the median RR and the number of GSR peaks. We have further investigated the joint use of these features with respect to their capability to cluster the data space. To this end, we partitioned SRAD data in two cluster using standard k-means algorithm as provided in MATLAB (Figure 19). Clustering was correlated with the dataset labels (either rest or driving/stress). This process led to an 87.9 % recognition rate for the rest state and a 92.3 % recognition rate for the driving state. The overall classification accuracy was found to be 89.4 %.

It is worth mentioning that using the same cluster centroids for MCA data we found a 90 % rate of correct classification (Figure 20). By taking a closer look at these results, we found that 100 % classification accuracy was not achieved as the algorithm failed to recognize the activation state of the two non-naïve subjects. Actually, they have had significant previous experience with the Stroop colour word test.

To better understand what could have been the contribution of additional features, we repeated the same process using other features. However, performances of k-means clustering deviated significantly from rest/activation data labelling.

Finally, we wish to point out that k-means was repeated 1000 times with random clustering initialization: we observed changes in final centroid position in less than 15 % of cases. However, even in these cases, displacement of cluster centroids was quite limited. Indeed, we observed a change in coordinates of  $< 1\%$  the total range of the feature space. Therefore, the final cluster centres were considered reliable.

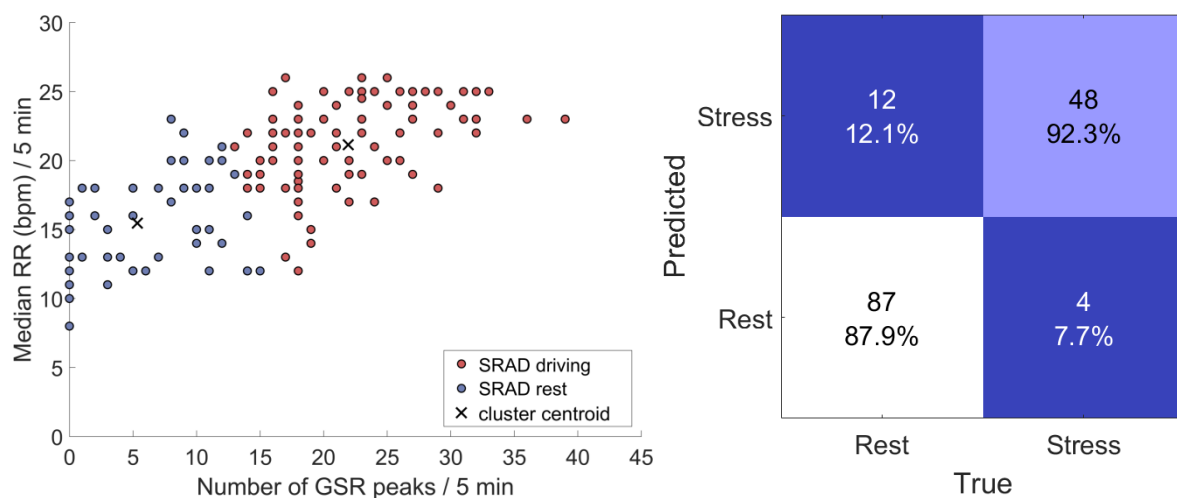


Figure 19 Classification performance for the SRAD dataset obtained using the first two features selected by the SFFS method. On the left a scatter plot of the dataset in the feature space, indicating the two clusters and related centroids. On the right the validation of the clusterization process.

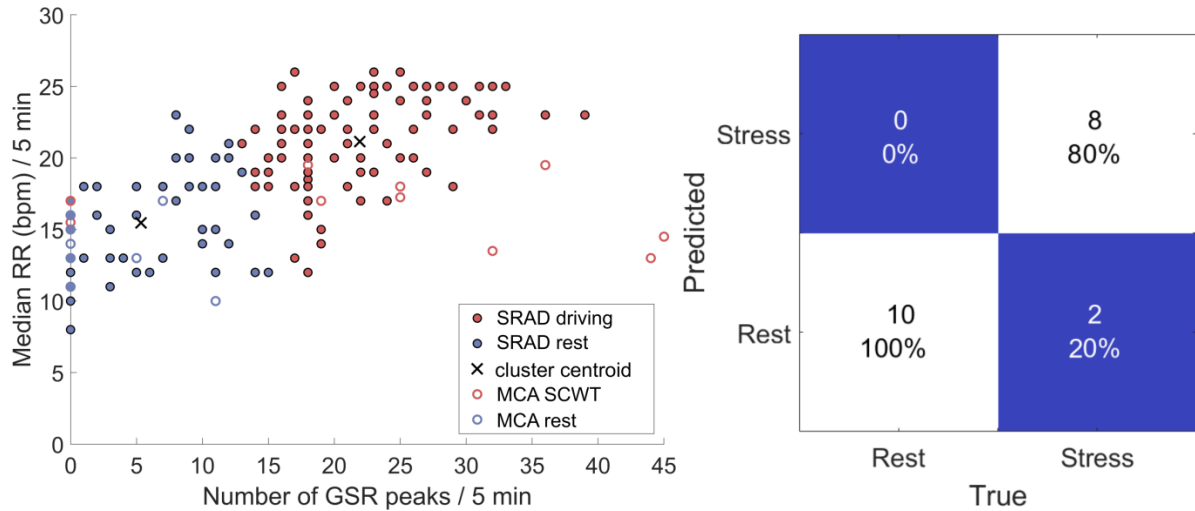


Figure 20 Classification performance for the MCA dataset obtained using the first two features selected by the SFFS method. On the left a scatter plot of the dataset in the feature space, indicating the two clusters and related centroids. On the right the validation of the clusterization process.

#### 7.4.4. Self-Organizing Maps

We have trained a set of two-dimensional SOMs with a number of units varying from 3 x 3 to 6 x 6. We did not consider larger maps due to the limited dataset size. Maps were trained on the SRAD dataset. For each map size, training was run ten times with random weight initialization. Apart changes in map orientation, no relevant differences were detected inside each run.

As we are interested in the topographic representation produced by SOMs, we have analysed each map with respect to the distance between weights of neighbouring units (the so called U-map), and the distribution in the network space of each weight dimension (weight-plane maps). In addition, to explore the semantic role of unit activation we analysed the distribution of data labels in network space (categorical hit maps). Since the results do not vary significantly with the number of units, to ease readability, we show only data for 4 x 4 and 5 x 5 SOMs in the mentioned figures.

Map distances in Figure 21 suggest that the units in the right upper corner are rather apart from the other units that tend to be closer each other. This confirms previous data from feature selection and data clustering, and also suggests that the data space can be partitioned into two highly structured clusters. It is worth noting that larger maps are expected to capture finer structural details of data as suggested by the comparison of the maps.

The distribution of SOM weights (Figure 22) provides additional support to the distance maps. In particular, spatial arrangement of weights looks consistent among different map sizes. In the maps of weight components (Figure 22) are represented units that respond to high or small values of the parameters. In addition, several components (e.g. those corresponding to median RR, number of GSR peaks, LF/HF, NN, LF, and HF) exhibit a well-

defined spatial distribution. In particular, some of them such as the weights of median RR and number of GSR peaks can be related to the partition appearing the left upper part of the map.

Categorical hit maps (Figure 23) for the rest and activation labels of the SRAD dataset shows a rather neat distinction among the two categories: only a few units respond simultaneously to both rest and activation data. A similar map is shown in Figure 24 for the MCA dataset; worth of note is that even this map displays two anomalous units corresponding to the non-naïve subjects. To summarize, results support that SOM has learned a topographic representation of the input space congruous with a priori data labels.

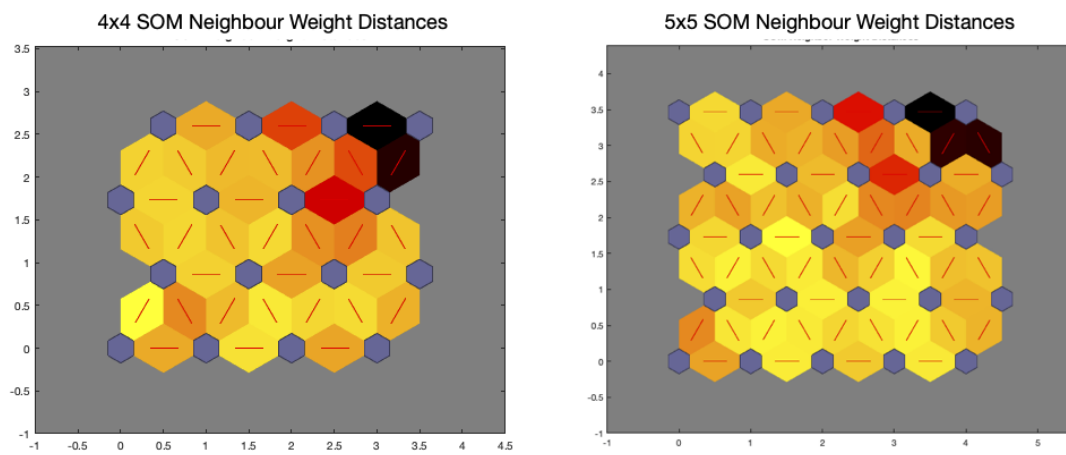


Figure 21 Maps of distances between the weight of neighboring units for 4 x 4 and 5 x 5 maps. Darkest colours indicate largest distance while light colours denotes smallest ones.

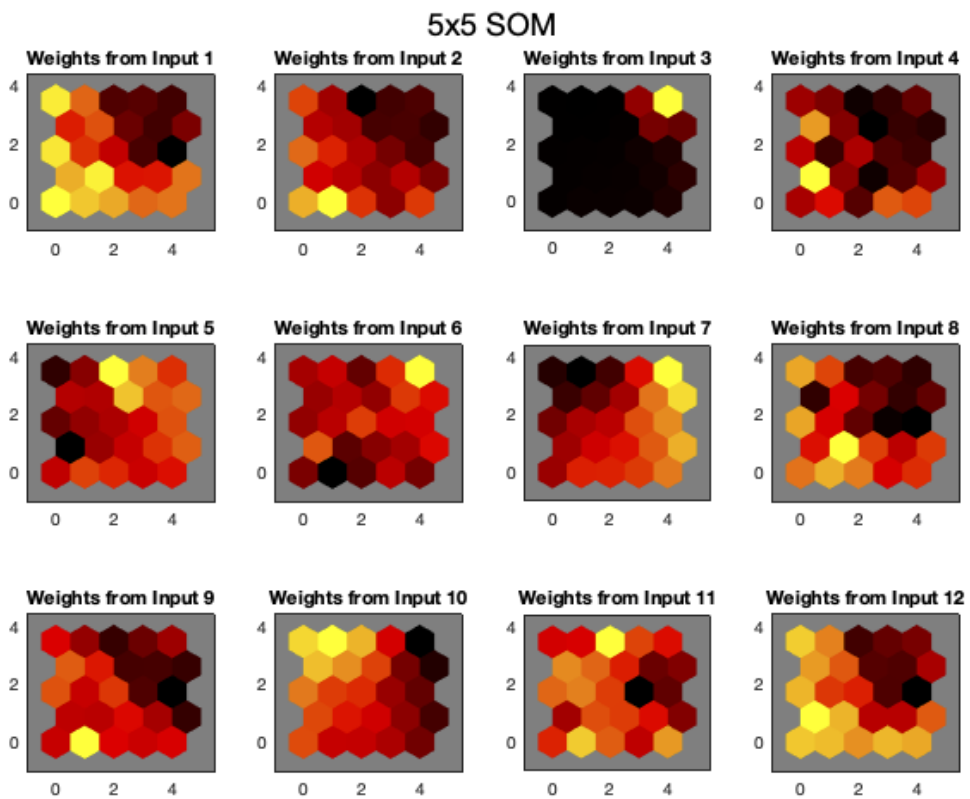
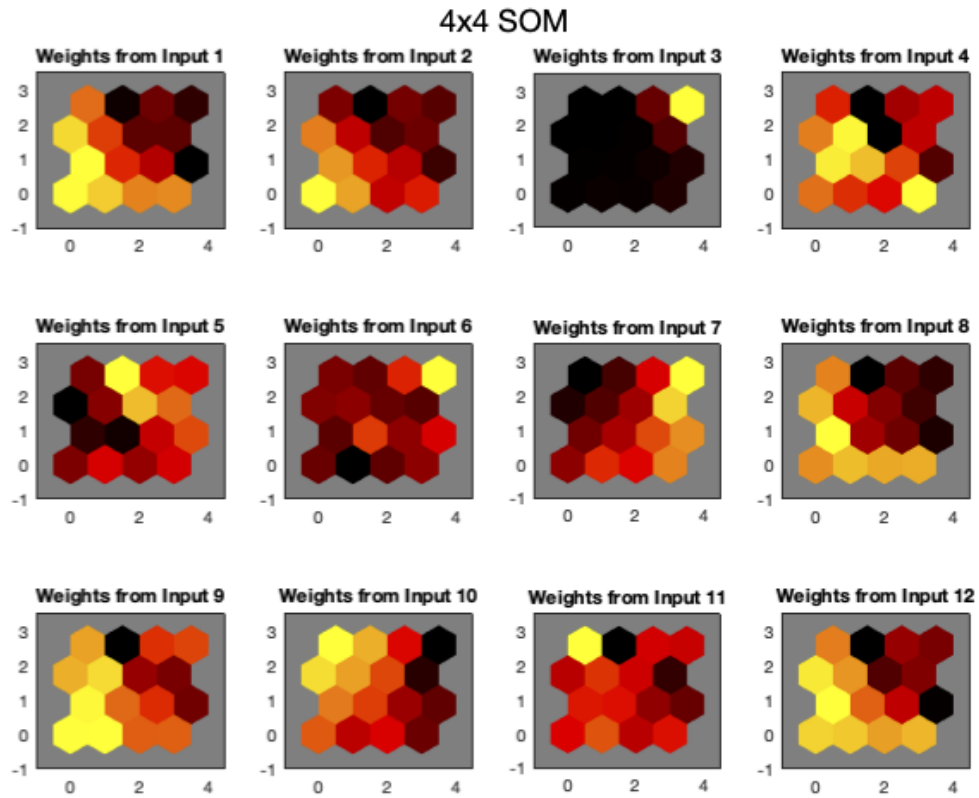


Figure 22 Maps of weight components. Darkest colors indicate smallest values while light colors denotes largest ones.

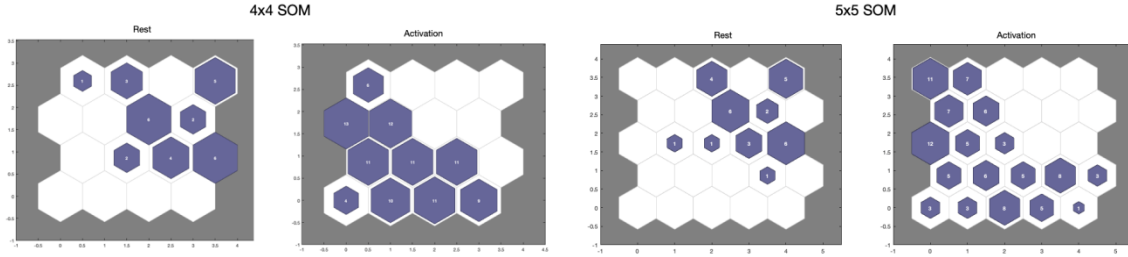


Figure 23 Map of winning units for SRAD data according to rest and activation labels.

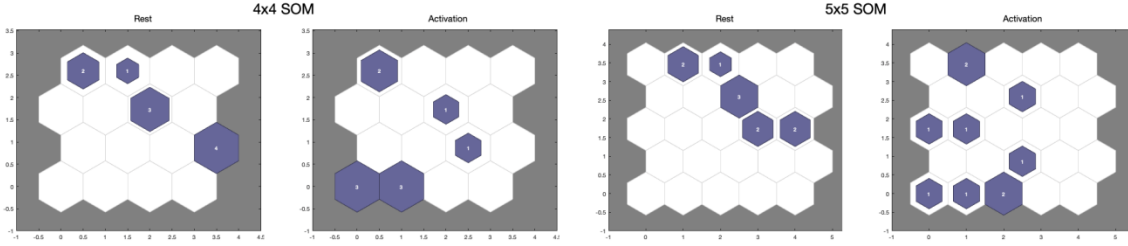


Figure 24 Map of winning units for MCA data according to rest and activation labels.

## 7.5. Discussion

In this chapter we enriched/integrated the “Mild Cognitive Activation” dataset collected from the volunteers enrolled during my work through the “Stress Recognition in Automobile Drivers” dataset from MIT Media lab. Aiming to assess stress activation, both of them were produced by recording a set of physiological signals in different settings. Our investigation was mainly conducted using the SRAD dataset (the most numerous) as development set, whilst MCA data were used for independent testing.

From the original data space we extracted a set of 12 features including descriptors of HR, HRV, RR and GSR which are sensitive to individual response to stressors with emphasis on ANS response. Analysis of SRAD features space by SFFS supports the conclusion that median RR and number of GSR peaks have a prominent discriminating power and can lead to recognize activation statuses, which is also confirmed by MCA data analysis.

We also tested AE features obtained from the SRAD dataset. They are estimated by using whole original data and are expected to reduce the potential information loss of SFFS mechanism. Results suggest that using two AE features can lead to good discrimination of rest states from activation ones. A similar conclusion is obtained by using the same AE features for MCA dataset. It stands to reason that our data space is intrinsically two-dimensional with respect recognition of activation condition. This conclusion is also in accord with SFFS results.

When applying standard k-means algorithm using the two most relevant features we observed two clusters that well represent rest and activation labels. This clustering is

consistent with MCA data. Interestingly, in the latter case, the activation state of two non-naïve subjects (the ones that had previous experience with the Stroop colour word test) was not detected. This actually confirms that the two subjects somehow behaved differently from the other ones. One could actually argue that the classification process did not really fail in this case, as activation of these subjects was quite moderate.

As to the final part of the work, it must be pointed out that individual responses are intrinsically variable and the use of flexible but compact representations of individual status are highly desirable. In this context, the use of SOM networks revealed promising.

Being non-supervised, SOM has the ability to autonomously discover significant pieces of information embedded in data space. In addition, they map data manifold in network space preserving topology and major related structural properties such as clustering.

SOMs derived on SRAD data show the existence of two virtually separated zones in the map: one of them tends to respond to rest statuses, whilst the other best matches activation statuses. These results are also confirmed by MCA data, not used to train the map. A significant aspect of SOMs that is relevant for applications is their ability to discover and represent the internal structure of large clusters. In particular, each map unit can be viewed as a prototype (or code) of the individual's status. In this view, activation (or rest) of a person is naturally represented by a structured family of codes.

# Conclusions

In this work we have designed and developed a measurement setup that allowed us the unobtrusive recording of basic physiological signals (Chapter 2, targets 1-3). The data so collected were used to successfully recognize stress conditions under mild cognitive activation.

A special attention was paid on the unobtrusive implementation of an effective processing method for measuring and analyzing the heart rate and its variability by using a video camera (Chapter 3). In this view we proposed the integration of multiple blood volume pulse signals, simultaneously acquired from a person's face. The method allows the reliable measurement of heart rate and its variability. In addition, respiratory rate, galvanic skin response, and EEG signals were analysed through a chest belt, hand sensors, and a headband, respectively.

As described in Chapter 4, Chapter 5, and Chapter 6 the results obtained from "Mild Cognitive Activation" dataset showed a good agreement with reports in scientific literature. This supports i) the robustness and reliability of our experimental setup for the unobtrusive monitoring, and ii) that the analysed signals are robust indicators of general stress conditions. The results obtained at this moment allowed the achievement of the target number four of the thesis.

The last target of the thesis was addressed in Chapter 7 and for this reason the dataset collected from the volunteers enrolled in this work was enriched/integrated with the "Stress Recognition in Automobile Drivers" dataset from the MIT Media Lab. A set of features were extracted from the signals available (including ECG, HR, RR, GSR, EMG acquired both at rest and during driving). The data harmonization and feature selection were performed using two complementary approaches: sequential forward feature selection and auto-encoder neural network. From an initial set of 12 features we have determined, by the analysis of feature space by SFFS, that 2 of them (median RR and number of GSR peaks) can discriminate activation statuses from resting ones; we have obtained similar results using the AE approach.

Then a cluster analysis was carried out using two complementary methods: k-means clustering and Kohonen self-organizing map. The preliminary unsupervised clustering carried out through the k-means clustering of the "Stress and Recognition in Automobile Drivers" exhibited interesting properties, enabling and efficient recognition of stress state. This result was further validated using the data from the MCA dataset. Also in this case, we were able to achieve good (90 %) discrimination accuracy.



Finally we explored the use of Kohonen self-organizing map to provide a flexible representation of an individual status. Finally we showed that SOM can provide a comprehensive but compact description of activation statuses.

Though the obtained results are in line with background literature, they are in a sense surprising as the two datasets used for development and testing were acquired under completely different experimental conditions (while driving or while performing the Stroop colour word test). This supports the idea that the used feature set is highly descriptive of individual activation status and able to predict a wide spectrum of activation conditions.

This can support the idea that our approach could generalize across experimental conditions and tasks, being a valid procedure to adopt for estimating personal wellness during daily activities.

# Appendix A

“Mild Cognitive Activation” Dataset – User provided data

Besides sensed data, during experiments each volunteer is asked to provide some basic personal details and fill short questionnaires pertaining her/his perceptions.

**At the beginning of the experiment**, the volunteer was asked to provide: age, sex, weight, height, and fill in the questionnaire of Perceived Stress Scale [127].

**At the end of each test phase**, the volunteers fills a short questionnaire about the stress level experienced by the subject. Stress self-assessment was conducted using two different questions as reported by Giokoumis et al. [128]. The first one is a Likert-scaled (1–5) question directly asking subjects whether they were feeling stressed during the test [129]. The second was a subset of the Stress Appraisal Measure questionnaire [130] including questions 2, 16, 24, and 26.

## **Appendix B**

The experimental protocol received the Ethical Clearance certification (0050349/2019, July 09th, 2019) by the Council National Research Committee for Research Ethics and Bioethics. In addition to that, written informed consent was obtained from all volunteers included in this study. In order to explain the experiment, an information sheet was also distributed to the volunteers.

**Istituto di Fisiologia Clinica del CNR**  
**Progetto: Quantitative Well-being Imaging Technology (QWITH)**

**Nota informativa al volontario in tutela della riservatezza dei propri dati personali**  
**Protezione dei dati personali**

Ogni informazione, dato personale che Lei riguardi ed il cui trattamento risulti connesso e indispensabile alla Sua partecipazione al presente studio, sarà trattato dal promotore con modalità idonee a garantire l'assoluta riservatezza, confidenzialità e sicurezza degli stessi, in conformità alle norme che regolano l'integrità della ricerca nonché a quelle per la tutela delle persone e di altri soggetti rispetto al trattamento di dati personali (D.Lgs. 30 giugno 2003, n. 196 (Codice Privacy) e successive modifiche e integrazioni, ed del Regolamento UE n. 2016/679 - "GDPR 2016/679"), Le chiediamo di tener conto di quanto nel seguito esplicitato.

**a) Finalità e modalità dei trattamenti**

I Suoi dati personali, oggetto di trattamento da parte del promotore di questo studio saranno costituiti dai dati comuni, le generalità (cognome e nome), che proprio in connessione alle esigenze di assoluta riservatezza, confidenzialità e sicurezza sopra menzionate, saranno criptate nel database dello studio. Detti dati raccolti presso di Lei o per Suo tramite dallo sperimentatore, su incarico del promotore, saranno da quest'ultimo registrati, elaborati, gestiti e archiviati – in forma cartacea, automatizzata e/o informatizzata – per le esclusive finalità connesse all'espletamento delle varie fasi del presente studio e, in particolare, al fine di verificare lo stato di avanzamento dello stesso. I dati personali saranno successivamente soggetti ad elaborazione statistica, e quindi, trasformati in forma totalmente anonima e, in questa forma, eventualmente inseriti in pubblicazioni e/o presentati in congressi, convegni e seminari a carattere scientifico.

**b) Natura obbligatoria o facoltativa dei dati e conseguenze di un eventuale rifiuto a fornire i dati**

Il consenso al trattamento dei suoi dati personali come sopra descritti, pur avendo natura facoltativa, risulta indispensabile ai fini dell'espletamento del presente studio nonché per l'adempimento dei connessi obblighi di legge. In assenza di detto conferimento, il promotore non potrà effettuare lo studio con la Sua partecipazione.

**c) Ambito di comunicazione e di diffusione**

I dati personali non saranno resi accessibili e disponibili a terzi. I dati raccolti non saranno inviati in paesi esteri e in particolare in paesi extra-comunitari che non adottino normative per garantire un livello di protezione dei dati personali pari a quanto previsto dalle leggi italiane. L'eventuale diffusione dei dati, per il tramite di pubblicazioni scientifiche e/o di presentazione in congressi, convegni e seminari, avverrà esclusivamente a seguito di un'elaborazione meramente statistica degli stessi e, quindi, in forma assolutamente anonima.

**d) Diritti dell'interessato ai sensi dell'art. 7 del D.Lgs. 196/2003 e degli articoli dal 15 al 22 del Regolamento UE n. 2016/679**

In qualità di interessato al trattamento dei dati personali (come definito dall'art. 4, comma 1, lettera i), Lei potrà in qualunque momento avvalersi della facoltà e dei diritti a Lei attribuiti ai sensi 7 del D.Lgs. 196/2003 e degli articoli dal 15 al 22 del Regolamento UE n. 2016/679 e più precisamente:

- potrà accedere al Registro generale dei trattamenti gestito dal Garante per la protezione dei dati personali;
- potrà essere informato in merito a quanto concerne:
  1. l'origine dei dati personali;
  2. le finalità e le modalità del trattamento;
  3. la logica applicata in caso di trattamento effettuato con l'ausilio di strumenti

- elettronici;
4. gli estremi identificativi del titolare, dei responsabili e del rappresentante designato ai sensi dell'articolo 5, comma 2 del D.Lgs. n. 196/2003;
  5. i soggetti o le categorie di soggetti ai quali i dati personali possono essere comunicati o che possono venirne a conoscenza in qualità di rappresentante designato nel territorio dello Stato, di responsabili incaricati.
- potrà ottenere, a cura del Titolare o del Responsabile del trattamento, senza ritardo:
    1. la conferma dell'esistenza e la comunicazione in forma comprensibile di dati personali che La riguardano, anche se non ancora registrati, e la loro comunicazione in forma intelligibile;
    2. la cancellazione, la trasformazione in forma anonima o il blocco dei dati trattati in violazione della legge, compresi quelli di cui non è necessaria la conservazione in relazione agli scopi per i quali i dati sono stati raccolti o successivamente trattati;
    3. l'aggiornamento, la rettifica e, quando vi ha interesse, l'integrazione dei dati;
    4. l'attestazione che le operazioni di cui numeri 2 e 3 sono state portate a conoscenza anche per quanto riguarda il loro contenuto, di coloro ai quali i dati sono stati comunicati o diffusi, eccettuato il caso in cui tale adempimento si rivela impossibile o comporta un impiego di mezzi manifestamente sproporzionato rispetto al diritto tutelato;
  - potrà opporsi, in tutto o in parte:
    1. per motivi legittimi al trattamento dei dati personali che la riguardano, ancorché pertinenti allo scopo della raccolta;
    2. al trattamento di dati personali che la riguardano a fini di invio di materiale pubblicitario o di vendita diretta o per il compimento di ricerche di mercato o di comunicazione commerciale.

#### **e) Titolare e Responsabile**

Il Titolare del trattamento dei Suoi dati personali è il Consiglio Nazionale delle Ricerche (CNR), promotore di questo studio presso la sede dell'Istituto di Fisiologia Clinica di Pisa.

Responsabile del trattamento è il responsabile dello studio, Dott. Giuseppe Coppini, domiciliato per la funzione presso il promotore (Istituto di Fisiologia Clinica di Pisa), in relazione alla attività di gestione dei dati.

**Istituto di Fisiologia Clinica del CNR**  
**Progetto: Quantitative Well-being Imaging Technology (QWITH)**

**Consenso al trattamento dei propri dati personali**

Io sottoscritto/a (nome e cognome in stampatello) .....

**DICHIARO**

- di aver preso visione della nota informativa al volontario in tutela della riservatezza dei propri dati personali e di aver compreso le informazioni in essa contenute e le informazioni integrative fornite in forma orale dal ricercatore
- di essere stato posto a conoscenza per iscritto e oralmente dei diritti esercitabili ai sensi dell'art. 7 del D.Lgs. n. 196/2003 (Codice Privacy) e degli articoli dal 15 al 22 del Regolamento UE n. 2016/679 (GDPR 2016/679)

Ai sensi dell' art 23 del D. Lgs. n. 196/2003, (Codice Privacy) e all'art. 13 del Regolamento UE n. 679/16 (GDPR 2016/679),

**AUTORIZZO**

il promotore Istituto di Fisiologia Clinica del CNR a sottoporre a trattamento in quanto necessari alla mia partecipazione allo studio, come specificato nel Foglio Informativo dello studio, dati personali che mi riguardano, raccolti, per incarico del promotore, dallo sperimentatore, per le finalità, nelle forme e nei modi specificamente descritti nella nota informativa allegata.

In particolare autorizzo il promotore a (*barrare la casella*):

☐ Trattare, oltre ai dati comuni, anche i miei dati personali così detti sensibili (cioè idonei a rilevare il mio stato di salute).

Firma \_\_\_\_\_ Data \_\_\_\_\_

# **Istituto di Fisiologia Clinica del CNR**

## **Progetto: Quantitative Well-being Imaging Technology (QWITH)**

### **Foglio informativo**

Gentile Signora/Gentile Signore,

Le viene chiesto di partecipare a una sperimentazione a scopo di ricerca. Le sperimentazioni includono solamente soggetti che scelgono volontariamente di parteciparvi.

Questo foglio ha lo scopo di fornire specifiche informazioni per aiutarLa a decidere se accettare o meno di partecipare. Il ricercatore responsabile dello studio è a Sua disposizione per rispondere a qualunque domanda. La preghiamo di leggere o ascoltare attentamente il contenuto di questo documento e di decidere, in assoluta libertà, se partecipare o meno.

#### **Introduzione**

Secondo l'Organizzazione Mondiale della Sanità, la salute dovrebbe essere definita come uno stato di completo benessere fisico, mentale e sociale e non meramente come l'assenza di malattia o infermità. Negli ultimi anni è cresciuta enormemente la consapevolezza circa l'importanza della prevenzione delle malattie con una conseguente rivalutazione del ruolo del benessere. Questo perché prevenzione è un insieme di attività, azioni ed interventi attuati con il fine prioritario di promuovere e conservare lo stato di salute ed evitare l'insorgenza di malattie. In particolar modo, lo stile di vita corretto aiuta a prevenire un vasto gruppo di malattie indicate spesso come *Non Communicable Diseases* (NCD), che include le cardiache, l'ictus, il cancro, il diabete e le malattie polmonari croniche. In tutto il mondo, le NCD uccidono più persone delle malattie infettive, sono responsabili di due terzi dei decessi ed hanno un notevole impatto socio-economico. Alcuni dei fattori che influenzano il rischio delle NCD sono in relazione allo stile di vita delle persone; tra i più significativi si ricordano alimentazione, attività fisica, uso di tabacco e alcool, stress e condizioni psicologiche avverse. D'altra parte si rileva una crescente consapevolezza individuale dell'importanza di uno stile di vita corretto ed una tendenza a diventare parte attiva nell'acquisizione e nel mantenimento del proprio stato di benessere. I sistemi per l'auto-monitoraggio e l'auto-apprendimento sono strumenti importanti per fornire alle persone la conoscenza ed i servizi per la cura del proprio benessere. Già da diversi anni sono stati sviluppati metodi e dispositivi per la valutazione dello stato di salute e di benessere dell'individuo basati sulla misurazione di parametri fisiologici, la valutazione dello stato psicologico e di parametri sociali.

#### **Descrizione dello studio**

La studio a cui Lei è invitata a partecipare è parte del progetto QWITH dell'Istituto di Fisiologia Clinica de CNR, che si propone di:

- individuare un insieme di parametri psicofisici indicativi dello stato di benessere dell'individuo, valutabili mediante dispositivi e metodi non intrusivi,
- realizzare un modello di integrazione di tali parametri al fine di produrre una descrizione quantitativa, utilizzabile dall'individuo per migliorare/mantenere il proprio stato di benessere.

A tal fine si prevede l'acquisizione di un insieme di segnali di riferimento da un gruppo di volontari.

Verranno in particolare misurati: frequenza cardiaca, saturazione di ossigeno, frequenza respiratoria, conduttanza della pelle, pressione arteriosa, movimenti oculari e segnale EEG. Le misure verranno effettuate sia in condizioni basali (assenza di stimoli specifici) che in presenza di stimoli durante l'esecuzione di un semplice test.

I segnali verranno acquisiti mediante telecamere, sensori di conduttanza della pelle, un dispositivo per misurare i movimenti oculari (eye-tracker), banda EEG, fascia toracica respiratoria, ECG, saturimetro e sfigmomanometro.

I dati verranno analizzati ed elaborati con tecniche di Intelligenza Artificiale per realizzare un modello quantitativo dello stato di benessere individuale.

Lo studio arruolerà 40 soggetti e si terrà interamente presso la sede di Pisa dell'Istituto di Fisiologia Clinica del CNR di Pisa. Tutte le attività dello studio verranno condotte sotto la supervisione medica della Dott.ssa Maria Aurora Morales che sarà a disposizione dei partecipanti per fornire i necessari chiarimenti ed illustrare il significato dei dati acquisiti.

### **Cosa Le viene chiesto**

Lei dovrà rispondere ad alcune domande fornendo alcuni dati anagrafici e fisiologici e completare alcuni questionari relativi alla sua attività fisica e alla percezione di un eventuale stato di stress. Tali informazioni verranno usate per interpretare ed integrare quanto osservato dai segnali acquisiti dai vari dispositivi.

Le verrà chiesto di indossare alcuni dispositivi (ovvero elettrodi ECG, banda toracica respiratoria, elettrodi per conduttanza della pelle e banda frontale EEG) e verranno acquisite immagini del suo volto tramite telecamere ed un eye tracker.

Questo ci consentirà di acquisire alcuni segnali (ovvero frequenza cardiaca, frequenza respiratoria, conduttanza della pelle, saturazione di ossigeno, pressione, informazioni sui movimenti oculari e segnale elettroencefalografico) sia in condizione basale che durante l'esecuzione dello *stroop color word test*. Si tratta di un semplice test della durata di circa 5 minuti, in cui Le verranno presentati in ordine casuale i nomi di alcuni colori (ad es. rosso, verde, giallo ...) visualizzati su un monitor con un colore diverso da quello del nome, ad ogni presentazione Le verrà chiesto di pronunciare il colore visualizzato.

### **Adesione allo studio**

La Sua partecipazione a questo studio è volontaria. Non dovrà partecipare a questo studio se non lo desidera. Ha il diritto di cambiare idea e di lasciare in qualunque momento lo studio senza dare alcuna motivazione, semplicemente informando uno dei ricercatori che conducono dell'esperimento.

Questo protocollo di ricerca nonché il contenuto di queste informazioni ed il modulo di Consenso Informato sono stati valutati ed approvati dalla Commissione per l'Etica della Ricerca e la Bioetica del CNR. Il responsabile del progetto Le chiederà di firmare e datare il modulo di Consenso Informato del presente documento prima che venga eseguita qualsiasi procedura o test per garantire che Lei sia stata/o informata/o in modo completo e che abbia dato liberamente il Suo consenso. L'originale del Consenso Informato scritto, firmato da Lei, verrà conservato presso l'archivio dell'Istituto di Fisiologia Clinica del CNR di Pisa mentre Le verrà consegnata una copia firmata di questo modulo di consenso da conservare.

Se ha qualunque domanda circa i suoi diritti come soggetto di ricerca lo studio potrà rivolgersi al responsabile del progetto.

il Dott. \_\_\_\_\_ Giuseppe Coppini \_\_\_\_\_ al numero \_ 050-3153480 \_



**Istituto di Fisiologia Clinica del CNR**  
**Progetto: Quantitative Well-being Imaging Technology (QWITH)**  
**Modulo di Consenso Informato**

Io sottoscritto/a .....

Dichiara di essere stato esaurientemente informato dalla lettura del Foglio Informativo e dallo sperimentatore riguardo allo studio.

Dichiaro inoltre che sono state date risposte esaustive e comprensibili alle domande da me formulate sempre riguardo a tutti i momenti della procedura sperimentale in esame.

Sono conscio del fatto che mi verrà fornito apposito modulo di consenso e l'informativa sulla privacy per la raccolta dei miei dati sensibili relativi allo studio in oggetto.

Io sottoscritto confermo, firmando questo modulo, di essere d'accordo a partecipare a questo studio e di aver ricevuto copia di questo modulo di Consenso Informato.

\_\_\_\_\_  
Firma del soggetto

\_\_\_\_\_  
Data

\_\_\_\_\_  
Firma dello sperimentatore

\_\_\_\_\_  
Data

# Bibliography

- [1] World Health Organization: Preamble to the Constitution of the World Health Organization, in: International Health Conference: 1946, New York, p. 100, 1946.
- [2] C. B. Corbin and R. P. Pangrazi, "Toward a Uniform Definition of Wellness: A Commentary", President's Council on Physical Fitness and Sports Research Digest, 2001.
- [3] R. Hubert, "Usability Field Study of Home Health Monitoring Devices Used by Older Adults", Proceedings of SRD2007, 2006.
- [4] W. C. Willett, "Balancing life-style and genomics research for disease prevention", Science, vol. 296, pp. 695–698, 2002.
- [5] J. Bousquet, J. M. Anto, P. J. Sterk et al. "Systems medicine and integrated care to combat chronic noncommunicable diseases", Genome Medicine, vol. 3, no. 7, pp. 1-12, 2011.
- [6] R. M. Kaplan, J. W. Bush, and C. C. Berry, "Health Status: Types of Validity and the Index of Well-being", Health Services Research, vol. 11, pp. 478-507, 1976.
- [7] C. L. Chiang, "An Index of Health: Mathematical Models", Public Health Service, 1965.
- [8] O. Steinbrocker, C. Traeger, and R. Battman, "Therapeutic Criteria in Rheumatoid Arthritis", J. Am. Med. Assoc., vol. 140, pp.765-770, 1949.
- [9] D. Karnofsky and J. Burchenal, "The Clinical Evaluation of Chemotherapeutic Agents in Cancer", In C. Macleod ed. Evaluation of Chemotherapeutic Agents, Columbia University Press, pp. 199-205, 1949.
- [10] The Criteria Committee of the New York Heart Association, Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels, pp. 253-256, 1994.
- [11] S. K. Bhavnani, C. K. Bichakjian, J. L. Schwartz, et al., "Getting patients to the right healthcare sources: From real-world questions to Strategy Hubs", Proceedings of AMIA' 2002, pp. 51-55, 2002.
- [12] International Wellbeing Group, "Personal Wellbeing Index: 5th Edition", Australian Centre on Quality of Life, Deakin University, 2013. Accessed online at <http://www.acqol.com.au/iwbg/wellbeing-index/index.php>.

- [13] J. Jr. Ware, M. Kosinski, and S. D. Keller, "A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity", *Med Care*, pp. 220-233, 1996.
- [14] L. M. Otto, A. Howerter, I. R. Bell, et al., "Exploring Measures of Whole Person Wellness: Integrative Well-Being and Psychological Flourishing", *Explore*, vol. 6, pp. 364-370, 2010.
- [15] R. Hubert, "Usability Field Study of Home Health Monitoring Devices Used by Older Adults", *Proceedings of SRD2007*, 2006.
- [16] Centers for Disease Control. "Overweight and obesity" Updated 2014. Accessed online at <http://www.cdc.gov/nchs/fastats/obesity-overweight.htm>.
- [17] A. Kailas, C.-C. Chong, and F. Watanabe, "From Mobile Phones to Personal Wellness Dashboards", *IEEE Pulse*, vol. 1, pp. 57-63, 2010.
- [18] A. Kailas, "A Generic Conceptual Model Linking Wellness, Health Lifestyles, and User Assistance", *IEEE 13th International Conference on e-Health Networking, Applications and Services*, pp. 266-269, 2011.
- [19] R.-Y. Huang and L.-R. Dung, "Measurement of heart rate variability using off-the-self smart phones", *Biomedical engineering online*, vol. 15, no. 11, 2016.
- [20] M. Frydrysiak, "Wearable Care System for Elderly People, *International Journal of Pharma Medicine and Biological Sciences*", vol. 5, no. 3, 2016.
- [21] H. Al-Libawy, A. Al-Ataby, W. Al-Nuaimy, et al., "HRV-Based Operator Fatigue Analysis and Classification Using Wearable Sensors", *Systems, Signals and Devices (SSD), 13th International Multi-Conference on. IEEE*, 2016.
- [22] Q. Ji and Z. Zhu, "Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue, *IEEE Transactions on Vehicular Technology*", vol. 53, no. 4, 2004.
- [23] I. R. Tayibnapis, D.Y. Koo, M. K. Choi, et al., "A Novel Driver Fatigue Monitoring using Optical Imaging of Face on Safe Driving System", *Control, Electronics, Renewable Energy and Communications (ICCEREC), International Conference on. IEEE*, 2016.
- [24] F. Laurent, M. Valderrama, M. Besserve, et al., "Multimodal information improves the rapid detection of mental fatigue", *Biomedical Signal Processing and Control*, vol. 8, no. 4, pp. 400-408, 2013.
- [25] F. Agrafioti, D. Hatzinakos, A. K. Anderson, et al., "ECG pattern analysis for emotion detection, *IEEE Transactions on Affective Computing*", vol. 3, no. 1, pp. 102-115, 2012.

- [26] H. Alemdar, and C. Ersoy, "Wireless sensor networks for healthcare: A survey", *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010. <https://doi.org/10.1016/j.comnet.2010.05.003>
- [27] M. J. Rodrigues, O. Postolache, and F. Cercas, "Physiological and behavior monitoring systems for smart healthcare environments: A review", *Sensors (Switzerland)*, vol. 20, no. 8, pp. 1–26, 2020. <https://doi.org/10.3390/s20082186>
- [28] H. Mshali, T. Lemlouma, M. Moloney, et al., "A survey on health monitoring systems for health smart homes", *International Journal of Industrial Ergonomics*, vol. 66, pp. 26–56, 2018. <https://doi.org/10.1016/j.ergon.2018.02.002>
- [29] T. Liang and Y. J. Yuan, "Wearable medical monitoring systems based on wireless networks: A review", *IEEE Sensors Journal*, vol. 16, no. 23, pp. 8186–8199, 2016. <https://doi.org/10.1109/JSEN.2016.2597312>
- [30] P. Henriquez, B. J. Matuszewski, Y. Andreu-Cabedo, et al., "Mirror mirror on the wall... an unobtrusive intelligent multisensory mirror for well-being status self-assessment and visualization", *IEEE Transactions on Multimedia*, 2017.
- [31] W. Liao, W. Zhang, Z. Zhu, et al., "Toward a decision-theoretic framework for affect recognition and user assistance", *International Journal of Human-Computer Studies*, vol. 64, no. 9, pp. 847-873, 2006.
- [32] Y. Tong, J. Chen, Q. Ji, et al., "A unified probabilistic framework for spontaneous facial action modeling and understanding", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 2, pp. 258-273, 2010.
- [33] W. Liao, W. Zhang, Z. Zhu, et al., "A real-time human stress monitoring system using dynamic Bayesian network", *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, pp. 70-70, 2005.
- [34] J.T. Cacioppo, L. Tassinary, and G. Berntson, "Handboook of psychophysiology", Cambridge university press, 2007.
- [35] H. Cohen, J. Benjamin, A. B. Geva, et al. "Autonomic dysregulation in panic disorder and in post-traumatic stress disorder: Application of power spectrum analysis of heart rate variability at rest and in response to recollection of trauma or panic attacks", *Psychiatry research*, vol. 96, no. 1, pp. 1–13, 2000.
- [36] J. W. Hughes and C. M. Stoney, "Depressed mood is related to high-frequency heart rate variability during stressors", *Psychosomatic medicine*, vol. 62, no. 6, pp. 796–803, 2000.

- [37] R. L.-C. Pan and J. K. J. Li, "A noninvasive parametric evaluation of stress effects on global cardiovascular function", *Cardiovascular Engineering*, 2007, vol. 7, no. 2, pp. 74–80, 2007.
- [38] S. Akselrod, D. Gordon, F.A. Ubel, et al. "Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control", *Science*, vol. 213, no. 4504, pp. 220–222, 1981.
- [39] H. Kurniawan, A. V. Maslov, and M. Pechenizkiy, "Stress detection from speech and galvanic skin response signals", *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pp. 209-214, 2013.
- [40] M. T. Cooney, E. Vartiainen, T. Laakitainen, et al., "Elevated resting heart rate is an independent risk factor for cardiovascular disease in healthy men and women", *Amer. Heart J.*, vol. 159, pp. 612–619.e3, 2010.
- [41] U. R. Acharya, K. P. Joseph, N. Kannathal, et al., "Heart rate variability: A review", *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, 2006.
- [42] E. Watanabe, K. Kiyono, Y. Yamamoto, et al., "Heart Rate Variability and Cardiac Diseases", Tokyo, Japan: Springer, pp. 163–178, 2017.
- [43] P. C. Dinas, Y. Koutedakis, and A. D. Flouris, "Effects of active and passive tobacco cigarette smoking on heart rate variability", *Int. J. Cardiology*, vol. 163, pp. 109–115, 2013.
- [44] B. S. Tegegne, T. Man, A. M. van Roon, et al., "Determinants of heart rate variability in the general population: The lifelines cohort study", *Heart Rhythm*, vol. 15, no. 10, pp. 1552–1558, 2018.
- [45] L. Capdevila, J. Moreno, J. Movellan, et al. "HRV based health&sport markers using video from the face", in *Proc. 34th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 5646–5649, 2012.
- [46] B.G. Lee, S. J. Jung, and W. Y. Chung, "Real-time physiological and vision monitoring of vehicle driver for non-intrusive drowsiness detection", *IET Commun.*, vol. 5, no. 17, pp. 2461–2469, Nov. 2011.
- [47] D. S. Quintana and J. A. J. Heathers, "Considerations in the assessment of heart rate variability in biobehavioral research", *Frontiers Psychol.*, vol. 5, pp. 1–10, 2014.
- [48] M. S. Buist, "Association between clinically abnormal observations and subsequent in-hospital mortality: a prospective study", *Resuscitation*, vol. 62, pp. 137-141, 2004.

- [49] T. J. Hodgetts, "The identification of risk factors for cardiac arrest and formulation of activation criteria to alert a medical emergency", *Resuscitation*, vol. 54, pp. 125-131, 2002.
- [50] M. Cretikos, "An evaluation of activation and implementation of the medical emergency team system", Sydney: The university of new south wales, 2006.
- [51] Y. Masaoka and I. Homma, "The effect of anticipatory anxiety on breathing and metabolism in humans", *Respiration physiology*, vol. 128, no. 2, pp. 171-177, 2001.
- [52] K. B. Nilsen, T. Sand, L. J. Stovner, et al. "Autonomic and muscular responses and recovery to one-hour laboratory mental stress in healthy subjects", *BMC Musculoskeletal Disorders*, vol. 8, no. 1, 2007.
- [53] R. F. Navea, P. J. Buenevenida, and C. D. Cruz, "Stress detection using Galvanic Skin Response: an android application", *Journal of Physics: conference series*, vol. 1372, no. 1, IOP Publishing, 2019.
- [54] J. W. Ahn, K. Yunseo, and C. K. Hee, "A novel wearable EEG and ECG recording system for stress assessment", *Sensors*, vol. 19, no. 9, 2019.
- [55] S.-H. Seo and J.-T. Lee, "Stress and EEG", In *ICHIT, InTech*, 2010. Available online: <https://www.intechopen.com/books/convergence-and-hybrid-information-technologies/stress-and-eeeg>.
- [56] J. Alonso, S. Romero, M. Ballester, et al., "Stress assessment based on EEG univariate features and functional connectivity measures", *Physiological measurement*, vol. 36, no. 7, pp. 1351, 2015.
- [57] A. F. Rabbi, A. Zony, P. De Leon, et al., "Mental workload and task engagement evaluation based on changes in electroencephalogram", *Biomedical Engineering Letters*, vol. 2, no. 3, pp. 139–146, 2012.
- [58] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion", *Brain Cogn.*, vol. 20, no. 1, pp. 125-151, 1992.
- [59] A. Bierhaus, J. Wolf, M. Addrassy, et al., "A mechanism converting psychosocial stress into mononuclear cell activation", *Proc. Natl. Acad. Sci. USA.*, vol. 100, no. 4, pp. 1920-1925, 2003.
- [60] E. S. Epel, E. H. Blackburn, J. Lin, et al., "Accelerated telomere shortening in response to life stress", *Proc. Natl. Acad. Sci. USA.*, vol. 101, no. 49, pp. 17312-17315, 2004.
- [61] S. C. Segerstrom and G.E. Miller, "Psychological stress and the human immune system: A meta-analytic study of 30 years of inquiry", *Psychol. Bull.*, vol. 130, no. 4, pp. 601-630, 2004.

- [62] S. Cohen, D. A. Tyrrell, and A. P. Smith, "Negative life events, perceived stress, negative affect, and susceptibility to the common cold", *J. Pers. Soc. Psychol.*, vol. 64, no. 1, pp. 131-140, 1993.
- [63] I. Papousek, G. and Schulter, "Covariations of EEG asymmetries and emotional states indicate that activity at frontopolar locations is particularly affected by state factors", *Psychophysiology*, vol. 39, no. 3, pp. 350-60, 2002.
- [64] G. Schulter and I. Papousek, "Bilateral electrodermal activity: relationships to state and trait characteristics of hemisphere asymmetry", *International Journal of Psychophysiology*, vol. 31, pp. 1-12, 1998.
- [65] S.-H. Seo, J.-T. Lee, and M. Crisan, "Stress and EEG", *Convergence and hybrid information technologies*, vol. 1, no. 1, pp. 413-424, 2010.
- [66] I. R. Tayibnapis, D. Y. Koo, M. K. Choi, et al. "A novel driver fatigue monitoring using optical imaging of face on safe driving system," *Proc. Int. Conf. Control, Electron., Renewable Energy Commun.*, pp. 115–120, 2016.
- [67] H. Monkaresi, N. Bosch, R. A. Calvo, et al. "Automated detection of engagement using video-based estimation of facial expressions and heart rate," *IEEE Trans. Affect. Comput.*, vol. 8, no. 1, pp. 15–28, Jan.–Mar. 2017.
- [68] A. Mesleh, D. Skopin, S. Baglikov, et al. "Heart rate extraction from vowel speech signals," *J. Comput. Sci. Technol.*, vol. 27, no. 6, pp. 1243–1251, 2012.
- [69] M. Garbey, N. Sun, A. Merla, et al. "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1418–1426, Aug. 2007.
- [70] O. Boric-Lubecke, V. M. Lubecke, I. Mostafanezhad, et al. "Doppler radar architectures and signal processing for heart rate extraction," *Mikrotalasna Revija*, vol. 15, no. 2, pp. 12–17, 2009.
- [71] D. Obeid, S. Sadek, G. Zaharia, et al., "Multitunable microwave system for touchless heartbeat detection and heart rate variability extraction," *Microw. Opt. Technol. Lett.*, vol. 52, no. 1, pp. 192–198, 2010.
- [72] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image," *Med. Eng. Phys.*, vol. 29, no. 8, pp. 853–857, 2007.
- [73] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.

- [74] F. Bousefsaf, C. Maaoui, and A. Pruski, "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 568–574, 2013.
- [75] T. Pursche, J. Krajewski, and R. Moeller, "Video-based heart rate measurement from human faces," *Proc. IEEE Int. Conf. Consum. Electron.*, pp. 544–545, 2012.
- [76] C. Brüser, C. H. Antink, T. Wartzek, et al. "Ambient and unobtrusive cardiorespiratory monitoring techniques," *IEEE Rev. Med. Eng.*, vol. 8, pp. 30–43, 2015.
- [77] H. Monkaresi, R. A. Calvo, and H. Yan, "A machine learning approach to improve contactless heart rate monitoring using a webcam," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 4, pp. 2168–2194, Jul. 2014.
- [78] L. Wei, Y. Tian, Y. Wang, et al. "Automatic webcam-based human heart rate measurements using Laplacian eigenmap," in *Proc. Asian Conf. Comput. Vision*, pp. 281–292, 2013.
- [79] L. Feng, L. M. Po, X. Xu, et al. "Motion-resistant remote imaging photoplethysmography based on the optical properties of skin," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 879–891, May 2015.
- [80] M. Lewandowska, J. Ruminski, T. Kocejko, et al. "Measuring pulse rate with a webcam—A non-contact method for evaluating cardiac activity," in *Proc. Federated Conf. Comput. Sci.*, pp. 405–410, 2011.
- [81] W. Wang, S. Stuijk, and G. De Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 415–425, Feb. 2015.
- [82] W. Verkruijsse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Express*, vol. 16, no. 26, pp. 21434–21445, 2008.
- [83] L. Tarassenko, M. Villarroel, A. Guazzi, et al. "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiological Meas.*, vol. 35, no. 5, pp. 807–31, 2014.
- [84] G. De Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Trans. Biomed. Eng.*, vol. 60, pp. 2878–2886, Oct. 2013.
- [85] F. Zhao, M. Li, Y. Qian, et al. "Remote measurements of heart and respiration rates for telemedicine," *PloS ONE*, vol. 8, Art. no. e71384. 2013.
- [86] W. Cui, L. E. Ostrander, and B. Y. Lee, "In vivo reflectance of blood and tissue as a function of light wavelength," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 632–639, Jun. 1990.



- [87] P. Düking, A. Hotho, H. C. Holmberg, et al., "Comparison of non-invasive individual monitoring of the training and health of athletes with commercially available wearable technologies", *Frontiers in physiology*, vol. 7, no. 71, 2016.
- [88] R. W. Picard and J. Scheirer, "The galvactivator: A glove that senses and communicates skin conductivity", *Proceedings 9th Int. Conf. on HCI*, 2001.
- [89] S. R. Vrana, "Emotional modulation of skin conductance and eyeblink responses to a startle probe", *Psychophysiology*, vol. 32, no. 4, pp. 351-357, 1995.
- [90] J. S. Lerner, R. E. Dahl, A. R. Hariri, et al., "Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses", *Biological psychiatry*, vol. 61, no. 2, pp. 253-260, 2006.
- [91] W. Liao, W. Zhang, Z. Zhu, et al., "A real-time human stress monitoring system using dynamic Bayesian network", In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, pp. 70-70, 2005.
- [92] F. Mokhayeri and M. R. Akbarzadeh-T, "Mental stress detection based on soft computing techniques", In *2011 IEEE International Conference on Bioinformatics and Biomedicine*, pp. 430-433, 2011.
- [93] J. R. Stroop, "Studies of interference in serial verbal reactions", *J. Exp. Psychol.* Vol. 18, pp. 643–662, 1935.
- [94] C. M. MacLeod and K. Dunbar, "Training and Stroop-like interference: evidence for a continuum of automaticity", *J. Exp. Psychol. Learn. Mem. Cogn.*, vol.14, pp. 126–135, 1988.
- [95] R. J. Ivnik, J. F. Malec, G. E. Smith, et al., "Neuropsychological tests' norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, STROOP, TMT, and JLO". *Clin. Neuropsychol*, vol. 10, pp. 262–278, 1996.
- [96] C. M. MacLeod, "Half a century of research on the Stroop effect: an integrative review", *Psychological bulletin*, vol. 109, no. 2, pp. 163, 1991.
- [97] D. Giakoumis, A. Drosou, P. Cipresso, et al., "Using activity-related behavioural features towards more effective automatic stress detection", *PloS one*, vol. 7, no. 9, e43571, 2012.
- [98] P. Renaud and J. P. Blondin, "The stress of Stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed", *International Journal of Psychophysiology*, vol. 27, no. 2, pp. 87-97, 1997.

- [99] C. Kirschbaum, K. M. Pirke, and D. H. Hellhammer, "The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting", *Neuropsychobiology*, vol. 28, no. 1-2, pp. 76-81, 1993.
- [100] S. S. Dickerson and M. E. Kemeny, "Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research", *Psychological bulletin*, vol. 130, no. 3, pp. 355, 2004.
- [101] H. Kurniawan, A. V. Maslov, and M. Pechenizkiy, "Stress detection from speech and galvanic skin response signals" In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pp. 209-214, 2013.
- [102] K. Dedovic, R. Renwick, N. K. Mahani, et al., "The Montreal Imaging Stress Task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain", *Journal of Psychiatry and Neuroscience*, vol. 30, no. 5, pp. 319–325, 2005.
- [103] M. Merino, I. Gómez, and A. J. Molina, "EEG feature variations under stress situations", In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6700-6703, 2015.
- [104] G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video", in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, pp. 3430–3437, 2013.
- [105] Task Force of the European Society of Cardiology and the North American Society for Pacing and Electrophysiology, "Heart rate variability. Standards of measurement, physiological interpretation, and clinical use", *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, 1996.
- [106] L. Iozzia, L. Cerina, and L. Mainardi, "Relationships between heart-rate variability and pulse-rate variability obtained from video-PPG signal using ZCA", *Physiological Meas.*, vol. 37, pp. 1934–1944, 2016.
- [107] W. H. Lin, D. Wu, C. Li, et al. "Comparison of heart rate variability from PPG with that from ECG," *IFMBE Proc. Int. Conf. Health Inform.*, pp. 213–215, 2014.
- [108] N. Pinheiro, R. Couceiro, J. Henriques, et al. "Can PPG be used for HRV analysis?", in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 2945–2949, 2016.
- [109] SEMEOTICONS Project. <https://www.semeoticons.eu>. [Online]. Available: <https://www.semeoticons.eu>
- [110] W. Wang, A. C. Den Brinker, S. Stuijk, et al., "Algorithmic principles of remote-PPG", *IEEE Trans. Biomed. Eng.*, vol. 64, no. 99, pp. 1479–1491, Jul. 2017.

- [111] “Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)”, Official J. Eur. Union, vol. L119, pp. 1–88, 2016. [Online]. Available: <http://eurlex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC>
- [112] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm”, IEEE Trans. Biomed. Eng., vol. 32, no. 3, pp. 230–236, Mar. 1985.
- [113] J. Vila, F. Palacios, J. Presedo, et al. “Time-frequency analysis of heart-rate variability”, IEEE Eng. Med. Biol. Mag., vol. 16, no. 5, pp. 119–126, Sep./Oct. 1997.
- [114] G. D. Clifford and L. Tarassenko, “Quantifying errors in spectral estimates of HRV due to beat replacement and resampling”, IEEE Trans. Biomed. Eng., vol. 52, no. 4, pp. 630–638, Apr. 2005.
- [115] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, “Heart rate variability: a review”, Medical and biological engineering and computing, vol. 44, no. 12, pp. 1031-1051, 2006.
- [116] I. Cygankiewicz and W. Zareba, “Heart rate variability”, Handbook of Clinical Neurology, vol. 117, pp. 379-393, 2013.
- [117] A. Hyvarinen, “Fast and robust fixed-point algorithms for independent component analysis”, IEEE Trans. Neural Netw., no. 3, pp. 626–634, May 1999.
- [118] T. Lindeberg, “Scale-Space Theory in Computer Vision”, Boston, MA, USA: Springer, 1994.
- [119] X. Jiang, G. B. Bian, and Z. Tian. “Removal of artifacts from EEG signals: a review”, Sensors, vol. 19, no. 5, 2019.
- [120] J. A. Healey and R. W. Picard, “Detecting stress during real-world driving tasks using physiological sensors”, IEEE Transactions on intelligent transportation systems, vol. 6, no. 2, pp. 156-166, 2005.
- [121] J. T. Cacioppo and L. G. Tassinary, “Inferring psychological significance from physiological signals,” Am. Psychol., vol. 45, no. 1, pp. 16–28, Jan. 1990.
- [122] I. Goodfellow, Y. Bengio, and Courville, “A. Deep Learning”, MIT Press, 2016. <http://www.deeplearningbook.org>.
- [123] G. E. Hinton, and R. Salakhutdinov, “Reducing the Dimensionality of Data with Neural Networks”, Science, vol. 313, no. 5786, pp. 504 – 507, 2006.

- [124] M. Dash and H. Liu, "Feature selection for classification", *Intelligent data analysis*, vol. 1, no. 3, pp. 131-156, 1997.
- [125] C. M. Bishop, "Pattern Recognition and Machine Learning", Springer: Singapore, 2006.
- [126] T. Kohonen, "Self-Organizing Maps", Springer: Berlin, 2001. doi:10.1007/978-3-642-56927-2.
- [127] S. Cohen, T. Kamarck, and R. Mermelstein, "Perceived stress scale. Measuring stress: A guide for health and social scientists", vol. 10, no. 1-2, 1994.
- [128] D. Giokoumis, A. Drosou, P. Cipresso, et al. "Using activity-related behavioural features towards more effective automatic stress detection", *PloS one*, vol. 7, no. 9, 2012.
- [129] J. A. Healey and R. W. Picard, "Detecting stress during real-word driving tasks using physiological sensor", *IEEE Transactions on intelligent transportation systems*, vol. 6, no. 2, pp. 156-166, 2005.
- [130] E. J. Peacock and P. T. P. Wong, "The stress appraisal measure (SAM): a multidimensional approach to cognitive appraisal", *Stress medicine*, vol. 6, no. 3, pp. 227-236, 1990.